

# On the Conditional Distributions and the Efficient Simulations of Exponential Integrals of Gaussian Random Fields

Jingchen Liu and Gongjun Xu  
Columbia University

March 3, 2013

## Abstract

In this paper, we consider the extreme behavior of a Gaussian random field  $f(t)$  living on a compact set  $T$ . In particular, we are interested in tail events associated with the integral  $\int_T e^{f(t)} dt$ . We construct a (non-Gaussian) random field whose distribution can be explicitly stated. This field approximates the conditional Gaussian random field  $f$  (given that  $\int_T e^{f(t)} dt$  exceeds a large value) in total variation. Based on this approximation, we show that the tail event of  $\int_T e^{f(t)} dt$  is asymptotically equivalent to the tail event of  $\sup_T \gamma(t)$  where  $\gamma(t)$  is a Gaussian process and it is an affine function of  $f(t)$  and its derivative field. In addition to the asymptotic description of the conditional field, we construct an efficient Monte Carlo estimator that runs in polynomial time of  $\log b$  to compute the probability  $P(\int_T e^{f(t)} dt > b)$  with a prescribed relative accuracy.

## 1 Introduction

Consider a Gaussian random field  $\{f(t) : t \in T\}$  living on a  $d$ -dimensional domain  $T \subset R^d$  with zero mean and unit variance, that is, for every finite subset  $\{t_1, \dots, t_n\} \subset T$ ,  $(f(t_1), \dots, f(t_n))$  is a mean zero multivariate Gaussian random vector. Let  $\mu(t)$  be a (deterministic) function and  $\sigma \in (0, \infty)$  be a scale factor. Define

$$\mathcal{I}(T) \triangleq \int_T e^{\sigma f(t) + \mu(t)} dt. \quad (1)$$

In this paper, we develop a precise asymptotic description of the conditional distribution of  $f$  given that  $\mathcal{I}(T)$  exceeds a large value  $b$ , that is,  $P(\cdot | \mathcal{I}(T) > b)$ . In particular, we provide a tractable total variation approximation (in the sample path space) for such conditional random fields based on a change of measure technique. In addition to the asymptotic descriptions, we design efficient Monte Carlo estimators that run in polynomial time of  $\log b$  for computing the tail probabilities

$$v(b) = P(\mathcal{I}(T) > b) = P\left(\int_T e^{\sigma f(t) + \mu(t)} dt > b\right) \quad (2)$$

with a prescribed relative accuracy.

**Applications.** The integral of exponential functions of Gaussian random fields plays an important role in many probability models. We present a few such models. In spatial point process modeling, let  $\lambda(t)$  be the intensity of a Poisson point process on  $T$ , denoted by  $\{N(A) : A \subset T\}$ . In order to build in spatial dependence structure, the log-intensity is typically modeled as a Gaussian process, that is,  $\log \lambda(t) = f(t) + \mu(t)$  and then  $E[N(A) | \lambda(\cdot)] = \int_A e^{f(t) + \mu(t)} dt$ , where  $\mu(t)$  is the

mean function and  $f(t)$  is a zero-mean Gaussian process. For instance, [20] considers the time series setting in which  $T$  is a one dimensional interval,  $\mu(t)$  is modeled as the observed covariate process and  $f(t)$  is an autoregressive process; see [23, 19, 50, 21, 22] for more examples. Under this setting, one can show that  $P(N(T) > b) \sim P(\int_T e^{f(t)+\mu(t)} dt > b)$  as  $b \rightarrow \infty$  (see [36]).

In portfolio risk analysis, consider a portfolio of  $n$  assets  $S_1, \dots, S_n$ . The asset prices are usually modeled as log-normal random variables. That is, let  $X_i = \log S_i$  and  $(X_1, \dots, X_n)$  follows a multivariate normal distribution. The total portfolio value  $S = \sum_{i=1}^n S_i$  is the sum of dependent log-normal random variables (see [25, 4, 11, 29, 24]). [7] derives the tail asymptotics of  $S$  when  $n$  is a fixed number. This asymptotic approximation can also be obtained by a more general result in [28]. If one can represent each asset price by a Gaussian random field at one location, that is,  $X_i = f(t_i)$ , then as the portfolio size becomes large and the asset prices become more correlated, the unit share price of the portfolio admits the limit  $\lim_{n \rightarrow \infty} S/n = \int e^{f(t)} d\vartheta(t)$  for some positive measure  $\vartheta(t)$ . See [14, 36] for detailed discussions on the random field representations of large portfolios.

In option pricing, the asset price (as a function of time) is typically modeled as a geometric Brownian motion ([13, 38]), that is,  $S(t) = e^{W(t)}$ , where  $W(t)$  is a Brownian motion. Then the payoff of an Asian call option with strike price  $K$  is  $\max(\int_0^T e^{W(t)} dt - K, 0)$ . The probability  $P(\int_0^T e^{W(t)} dt > K)$  is the expected payoff of a digital Asian call option.

**The literature.** In the probability literature, the extreme behaviors of Gaussian random fields have been studied extensively. The results range from general bounds to sharp asymptotic approximations. An incomplete list of works includes [30, 32, 37, 42, 18, 47, 12, 33, 44]. A few lines of investigations on the supremum norm are given as follows. Assuming locally stationary structure, the double-sum method ([41]) provides the exact asymptotic approximation of  $\sup_T f(t)$  over a compact set  $T$ , which is allowed to grow as the threshold tends to infinity. For almost surely at least twice differentiable fields, [1, 45, 3] derive the analytic form of the expected Euler-Poincaré Characteristics of the excursion set  $(\chi(A_b))$  which serves as a good approximation of the tail probability of the supremum. The tube method ([43]) takes advantage of the Karhune-Loève expansion and Weyl's formula. A recent related work along this line is given by [40]. The Rice method ([8, 9, 10]) provides an implicit description of  $\sup_T f(t)$ . The discussions also go beyond the Gaussian fields. For instance, [31] discusses the situations of Gaussian process with random variances. See also [2] for discussions on non-Gaussian cases. The distribution of  $\mathcal{I}(T)$  is studied in the literature when  $f(t)$  is a Brownian motion ([49, 26]). Recently, [34, 36] derive the asymptotic approximations of  $P(\mathcal{I}(T) > b)$  as  $b \rightarrow \infty$  for three times differentiable and homogeneous Gaussian random fields.

Besides the tail probability approximations, rigorous analysis of the conditional distributions of stochastic processes given the occurrence of rare events is also an important topic. In the classic large deviations analysis for light-tailed stochastic systems, the sample path(s) that admits the highest probability (the most likely sample path) under the conditional distribution given the occurrence of a rare event is central to the entire analysis in terms of determining the appropriate exponential change of measure, developing approximations of the tail probabilities, and designing efficient simulation algorithms (see, for instance, standard textbook [27]). For heavy-tailed systems, the conditional distributions and the most likely paths, which typically admit the so-called “one-big-jump” principle, are also intensively studied ([5, 6, 17]). These results not only provide intuitive and qualitative descriptions of the conditional distribution but also shed light on the design of rare-event simulation algorithms ([15, 16, 17]) – the best importance sampling estimator of the rare-event probability uses a change of measure corresponding to the interesting conditional distribution. In addition, the conditional distribution (or the conditional expectations) is also of practical interest.

For instance, in risk management, the conditional expected loss given some rare/disastrous event is an important risk measure.

**Contributions.** In this paper, we pursue along this line for the extreme behaviors of Gaussian processes and begin to describe the conditional distribution of  $f$  given the occurrence of the event  $\{\mathcal{I}(T) > b\}$ . In particular, we provide both quantitative and qualitative descriptions of this conditional distribution. Furthermore, from a computational point of view, we construct a Monte Carlo estimator that takes a polynomial computational cost (in  $\log b$ ) to estimate  $v(b)$  for a prescribed relative accuracy.

Central to the analysis is the construction of a change of measure on the space  $C(T)$  (continuous functions living on  $T$ ). The application of the change of measure ideas is common in the study of large deviations analysis for the light-tailed stochastic systems. However, it is not at all standard in the study of Gaussian random fields. The proposed change of measure is not of a classical exponential-tilting form. This measure has several features that are appealing both theoretically and computationally. First, we show that the change of measure denoted by  $Q$  approximates the conditional measure  $P(\cdot|\mathcal{I}(T) > b)$  in total variation as  $b \rightarrow \infty$ . Second, the measure  $Q$  is analytically tractable in the sense that the distribution of  $f$  under  $Q$  has a closed form representation and the Randon-Nikodym derivative  $dQ/dP$  takes the form of a  $d$ -dimensional integral. This tractability property has useful consequences. From a methodological point of view, the measure  $Q$  provides a very precise description of the mechanism that drives the rare event  $\{\mathcal{I}(T) > b\}$ . This result allows to directly use the intuitive mechanism to provide functional probabilistic descriptions that emphasize the most important elements that are present in the interesting rare events. More technically, the analytical computations associated with the measure  $Q$  are easy (compared to the conditional measure) and the expectation  $E^Q[\cdot]$  is theoretically much more tractable than  $E[\cdot|\mathcal{I}(T) > b]$ . Based on this result, we show that the tail event  $\{\mathcal{I}(T) > b\}$  is asymptotically equivalent to the tail event of  $\sup_T \gamma(t)$  where  $\gamma(t)$  is an affine function of  $f(t)$  and its derivative field  $\partial^2 f(t)$  and  $\gamma(t)$  implicitly depends on  $b$ .

Another contribution of this paper lies in the numerical evaluation of  $v(b)$ . The importance sampling algorithm associated with the proposed change of measure yields an efficient estimator for computing  $v(b)$ . An important issue concerns the implementation of the Monte Carlo method. The processes considered in this paper are continuous while computers can only represent discrete objects. Inevitably, we will introduce a suitable discretization scheme and use discrete (random) objects to approximate the continuous processes. A naturally raised issue lies in the control of the approximation error relative to the probability  $v(b)$ . We will perform careful analysis and report the overall computational complexity of the proposed Monte Carlo estimators.

The rest of this paper is organized as follows. In Section 2, we present the main results including the change of measure, the approximation of  $P(\cdot|\mathcal{I}(T) > b)$ , and the efficient Monte Carlo estimator of  $v(b)$ . Proofs of the theorems are given in Sections 3-6. An appendix is added including all the supporting lemmas.

## 2 Main results

### 2.1 Problem setting and notations

Throughout the discussion, we consider a homogeneous Gaussian random field  $\{f(t) : t \in T\}$  living on a domain  $T \subset \mathbb{R}^d$ . Let the covariance function be

$$C(t-s) = \text{Cov}(f(t), f(s)).$$

We impose the following assumptions:

- C1  $f$  is homogenous with  $Ef(t) = 0$  and  $Ef^2(t) = 1$ .
- C2  $f$  is almost surely at least three times differentiable with respect to  $t$ .
- C3  $T$  is a  $d$ -dimensional Borel measurable compact set of  $R^d$  with piecewise smooth boundary.
- C4 The Hessian matrix of  $C(t)$  at the origin is standardized to be  $-I$ , where  $I$  is the  $d \times d$  identity matrix.
- C5 For each  $t \in R^d$ , the function  $C(\lambda t)$  is a non-increasing function of  $\lambda \in R^+$ .
- C6 The mean function  $\mu(t)$  is three-time differentiable. In addition, it falls into either of the two cases:  $\mu(t) \equiv 0$  or the maximum of  $\mu(t)$  is unique and is attained in the interior of  $T$ .

We define a set of notations constantly used in the later development and provide some basic calculations. Let  $P_b^*$  be the conditional measure given  $\{\mathcal{I}(T) > b\}$ , that is,

$$P_b^*(f(\cdot) \in A) = P(f(\cdot) \in A | \mathcal{I}(T) > b).$$

Let “ $\partial$ ” denote the gradient and “ $\Delta$ ” denote the Hessian matrix with respect to  $t$ . The notation “ $\partial^2$ ” is used to denote the vector of second derivatives. The difference between  $\partial^2 f(t)$  and  $\Delta f(t)$  is that  $\Delta f(t)$  is a  $d \times d$  symmetric matrix whose diagonal and upper triangle consist of elements of  $\partial^2 f(t)$ . Furthermore, let  $\partial_j f(t)$  be the partial derivative with respect to the  $j$ -th element of  $t$ . Lastly, we define the following vectors

$$\begin{aligned} \mu_1(t) &= -(\partial_1 C(t), \dots, \partial_d C(t)), \\ \mu_2(t) &= \left( \partial_{ii}^2 C(t), i = 1, \dots, d; \partial_{ij}^2 C(t), i = 1, \dots, d-1, j = i+1, \dots, d \right), \\ \mu_{02}^\top &= \mu_{20} = \mu_2(0). \end{aligned} \tag{3}$$

Suppose  $0 \in T$ . It is well known that (c.f. Chapter 5.5 of [3])  $(f(0), \partial^2 f(0), \partial f(0), f(t))$  is a multivariate Gaussian random vector with mean zero and covariance matrix

$$\begin{pmatrix} 1 & \mu_{20} & 0 & C(t) \\ \mu_{02} & \mu_{22} & 0 & \mu_2^\top(t) \\ 0 & 0 & I & \mu_1^\top(t) \\ C(t) & \mu_2(t) & \mu_1(t) & 1 \end{pmatrix}$$

where the matrix  $\mu_{22}$  is a  $d(d+1)/2$ -dimensional positive definite matrix and contains the 4th order spectral moments arranged in an appropriate order according to the order of elements in  $\partial^2 f(0)$ . Let  $h(x, y, z)$  be the density function of  $(f(t), \partial f(t), \partial^2 f(t))$  evaluated at  $(x, y, z)$ . Then, simple calculations yield that

$$h(x, y, z) = \frac{\det(\Gamma)^{-\frac{1}{2}}}{(2\pi)^{\frac{(d+1)(d+2)}{4}}} e^{-\frac{1}{2} \left[ y^\top y + \frac{(x - \mu_{20} \mu_{22}^{-1} z)^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + z^\top \mu_{22}^{-1} z \right]}, \tag{4}$$

where  $\det(\cdot)$  is the determinant of a matrix and

$$\Gamma = \begin{pmatrix} 1 & \mu_{20} \\ \mu_{02} & \mu_{22} \end{pmatrix}.$$

We define  $u$  as a function of  $b$  such that

$$\left(\frac{2\pi}{\sigma}\right)^{\frac{d}{2}} u^{-\frac{d}{2}} e^{\sigma u} = b. \quad (5)$$

Note that the above equation generally has two solutions, one is approximately  $\sigma^{-1} \log b$  and the other is close to zero. We choose  $u$  to be the one close to  $\sigma^{-1} \log b$ . For  $\mu(t)$  and  $\sigma$  appearing in (1), we define

$$\mu_\sigma(t) = \mu(t)/\sigma, \quad u_t = u - \mu_\sigma(t). \quad (6)$$

Approximately,  $u_t$  is the level that  $f(t)$  needs to reach so that  $\mathcal{I}(T) > b$ . Furthermore, we need the following spatially varying set:

$$A_t = \{f(\cdot) \in C(T) : \alpha_t > u_t - \eta u_t^{-1}\}, \quad (7)$$

where  $\eta > 0$  is a tuning parameter that will be eventually sent to zero as  $b \rightarrow \infty$  and  $\alpha_t$  is a function of  $f(t)$  and its derivative fields taking the form of

$$\alpha_t = f(t) + \frac{|\partial f(t)|^2}{2u_t} + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t}. \quad (8)$$

In the above equation (8),  $\bar{f}_t''$  is defined as (with the notations in (3))

$$\bar{f}_t'' = \partial^2 f(t) - u_t \mu_{02}. \quad (9)$$

The term  $B_t$  is a deterministic function depending only on  $C(t)$ ,  $\mu(t)$ , and  $\sigma$ ,

$$B_t = \frac{\mathbf{1}^\top \partial^2 \mu_\sigma(t) + d \times \mu_\sigma(t)}{2\sigma} + \frac{1}{8\sigma^2} \sum_i \partial_{iii}^4 C(0) + |\partial \mu_\sigma(t)|^2, \quad (10)$$

where  $d$  is the dimension of  $T$ , and  $\mathbf{1} = (\underbrace{1, \dots, 1}_d, \underbrace{0, \dots, 0}_{d(d-1)/2})^\top$ . Note that  $\alpha_t \approx f(t)$ . Thus, on the set

$A_t$ ,  $f(t) \approx \alpha_t > u_t - O(u^{-1})$ . Together with the fact that  $E[\partial^2 f(t) | f(t) = u_t] = u_t \mu_{02}$ ,  $\bar{f}_t''$  is the standardized second derivative of  $f$  given that  $f(t) = u_t$ . In Section 2.2, we will show that the event  $\{\mathcal{I}(T) > b\}$  is approximately  $\cup_{t \in T} A_t$ .

For notational convenience, we write  $a_u = O(b_u)$  if there exists a constant  $c > 0$  independent of everything such that  $a_u \leq cb_u$  for all  $u > 1$ , and  $a_u = o(b_u)$  if  $a_u/b_u \rightarrow 0$  as  $u \rightarrow \infty$  and the convergence is uniform in other quantities. We write  $a_u = \Theta(b_u)$  if  $a_u = O(b_u)$  and  $b_u = O(a_u)$ . In addition, we write  $a_u \sim b_u$  if  $a_u/b_u \rightarrow 1$  as  $u \rightarrow \infty$ .

**Remark 1** Condition C1 assumes unit variance. We treat the standard deviation  $\sigma$  as an additional parameter and consider  $\int e^{\mu(t) + \sigma f(t)} dt$ . Condition C2 is rather a strong assumption. It implies that  $C(t)$  is at least 6 times differentiable and the first, third, and fifth derivatives at the origin are all zero. Differentiability is a crucial assumption in this analysis. Condition C3 restricts the results to finite horizon. Condition C4 is introduced to simplify notations. For any Gaussian process  $g(t)$  with covariance function  $C_g(t)$  and  $\Delta C_g(0) = -\Sigma$  and  $\det(\Sigma) > 0$ , C4 can be obtained by an affine transformation by letting  $g(t) = f(\Sigma^{1/2}t)$  and

$$\int_T e^{\mu(t) + \sigma g(t)} dt = \det(\Sigma^{-1/2}) \int_{\{s: \Sigma^{-1/2}s \in T\}} e^{\mu(\Sigma^{-1/2}s) + \sigma f(s)} ds.$$

Conditions C5 is imposed for technical reasons so that we are able to localize the integration. For condition C6, we assume that  $\mu(t)$  is either a constant or attains its global maximum at one place. If  $\mu(t)$  has multiple (finitely many) maxima, the techniques developed in this paper still apply, but the derivations will be more tedious. Therefore, we stick to the uni-mode case.

**Remark 2** The setting in (2) also incorporates the case in which the integral is with respect to other measures with smooth densities. Then, if  $\nu(dt) = \kappa(t)dt$ , we will have that

$$\int_A e^{\mu(t)+\sigma f(t)} \nu(dt) = \int_A e^{\mu(t)+\log \kappa(t)+\sigma f(t)} dt,$$

which shows that the density can be absorbed by the mean function.

## 2.2 Approximation of the conditional distribution

In this subsection, we propose a change of measure  $Q$  on the sample path space  $C(T)$  that approximates  $P_b^*$  in total variation. Let  $P$  be the original measure. The measure  $Q$  is defined such that  $P$  and  $Q$  are mutually absolutely continuous. We define the measure  $Q$  under two different scenarios:  $\mu(t)$  is not a constant and  $\mu(t) \equiv 0$ . Note that the measure  $Q$  obviously will depend on  $b$ . To simplify the notations, we omit the index  $b$  in  $Q$  whenever there is no ambiguity.

The measure  $Q$  takes a mixture form of three measures, which are weighted by  $(1 - \rho_1 - \rho_2)$ ,  $\rho_1$ , and  $\rho_2$  respectively (a natural constraint is that  $\rho_1, \rho_2$ , and  $1 - \rho_1 - \rho_2 \in [0, 1]$ ). We define  $Q$  through the Radon–Nikodym derivative

$$\frac{dQ}{dP} = (1 - \rho_1 - \rho_2) \int_T l(t) \cdot LR(t) dt + \rho_1 \int_T l(t) \cdot LR_1(t) dt + \rho_2 \int_T \frac{LR_2(t)}{mes(T)} dt, \quad (11)$$

where  $\rho_1, \rho_2$  will be eventually sent to 0 as  $b$  goes to infinity,  $mes(T)$  is the Lebesgue measure of  $T$ , and

$$LR(t) = \frac{h_{0,t}(f(t), \partial f(t), \partial^2 f(t))}{h(f(t), \partial f(t), \partial^2 f(t))}, \quad LR_1(t) = \frac{h_{1,t}(f(t), \partial f(t), \partial^2 f(t))}{h(f(t), \partial f(t), \partial^2 f(t))}, \quad LR_2(t) = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(f(t)-u_t)^2}}{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}f(t)^2}}. \quad (12)$$

The density  $h(f(t), \partial f(t), \partial^2 f(t))$  is defined in (4),  $l(t)$  is a density function on  $T$ ,  $h_{0,t}$  and  $h_{1,t}$  are densities of  $(f(t), \partial f(t), \partial^2 f(t))$ . Before presenting the specific forms of  $l(t)$ ,  $h_{0,t}$ , and  $h_{1,t}$ , we would like to provide an intuitive explanation of  $dQ/dP$  from a simulation point of view. One can generate  $f(t)$  under the measure  $Q$  via the following steps:

1. Generate  $\iota \sim \text{Bernoulli}(\rho_2)$ .
2. If  $\iota = 1$ , then
  - (a) Generate  $\tau$  uniformly from the index set  $T$ , i.e.,  $\tau \sim \text{Unif}(T)$ .
  - (b) Given the realized  $\tau$ , generate  $f(\tau) \sim N(u_\tau, 1)$ .
  - (c) Given  $(\tau, f(\tau))$ , simulate  $\{f(t) : t \neq \tau\}$  from the original conditional distribution under  $P$ .
3. If  $\iota = 0$ 
  - (a) Simulate a random variable  $\tau$  following the density function  $l(t)$ .
  - (b) Given the realized  $\tau$ , simulate  $f(\tau) = x, \partial f(\tau) = y, \partial^2 f(\tau) = z$  from density function

$$h_{all}(x, y, z) = \frac{1 - \rho_1 - \rho_2}{1 - \rho_2} h_{0,\tau}(x, y, z) + \frac{\rho_1}{1 - \rho_2} h_{1,\tau}(x, y, z). \quad (13)$$

- (c) Given  $(\tau, f(\tau), \partial f(\tau), \partial^2 f(\tau))$ , simulate  $\{f(t) : t \neq \tau\}$  from the original conditional distribution under  $P$ .

Thus,  $\tau$  is a random index at which we twist the distribution of  $f$  and its derivatives. The likelihood ratio at a specific location  $\tau$  is given by  $LR(\tau)$ ,  $LR_1(\tau)$ , or  $LR_2(\tau)$  depending on the mixture component. The distribution of the rest of the field  $\{f(t) : t \neq \tau\}$  given  $(f(\tau), \partial f(\tau), \partial^2 f(\tau))$  is the same as that under  $P$ . It is not hard to verify that the above simulation procedure is consistent with the Radon-Nikodym derivative in (11).

We now provide the specific forms of the functions defining  $Q$ . We first consider the situation when  $\mu(t) \neq 0$ . By condition C6,  $\mu(t)$  admits its unique maximum at  $t_* = \arg \sup_{t \in T} \mu(t)$  in the interior of  $T$ . Furthermore, the Hessian matrix  $\Delta \mu_\sigma(t_*)$  is negative definite. The function  $l(t)$  is a density on  $T$  such that for  $t \in T$

$$l(t) = (1 + o(1)) \det(\Delta \mu_\sigma(t_*))^{1/2} \left( \frac{u_{t_*}}{2\pi} \right)^{d/2} e^{-\frac{u_{t_*}}{2} (t - t_*)^\top \Delta \mu_\sigma(t_*) (t - t_*)}, \quad (14)$$

which is approximately a Gaussian density centered around  $t_*$ . The functions  $h_{0,t}$  and  $h_{1,t}$  are density functions of  $(f(t), \partial f(t), \partial^2 f(t))$  defined as follows (we will explain the following complicated functions momentarily)

$$\begin{aligned} h_{0,t}(f(t), \partial f(t), \partial^2 f(t)) &= \mathbb{I}_{A_t} \times H_\lambda \times u_t \times e^{-\lambda u_t \left( f(t) + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} - u_t \right)} \times e^{-\frac{|\partial f(t)|^2}{2}} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} \bar{f}_t''|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} \bar{f}_t'' - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right] \right\}, \\ h_{1,t}(f(t), \partial f(t), \partial^2 f(t)) &= \mathbb{I}_{A_t^c} \times H_{\lambda_1} \times u_t \times e^{\lambda_1 u_t \left( f(t) + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} - u_t \right)} \times e^{-\frac{|\partial f(t)|^2}{2}} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} \bar{f}_t''|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} \bar{f}_t'' - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right] \right\}. \end{aligned}$$

where  $\mathbb{I}$  is the indicator function,  $A_t = \{f(\cdot) : f(t) + \frac{|\partial f(t)|^2}{2u_t} + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} > u_t - \eta/u_t\}$  is defined as in (7),  $\bar{f}_t''$  is defined as in (9),  $\lambda < 1$  is positive and it will be sent to 1 as  $b$  goes to infinity,  $\lambda_1$  is a fixed positive constant (e.g.,  $\lambda_1 = 1$ ), and the normalizing constants are defined as

$$\begin{aligned} H_\lambda &= \frac{e^{-\lambda \eta} (1 - \lambda)^{d/2} \lambda}{(2\pi)^{\frac{d}{2}}} \times \left[ \int_{z \in R^{d(d+1)/2}} e^{-\frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} z|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} z - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right]} dz \right]^{-1}, \\ H_{\lambda_1} &= \frac{e^{\lambda_1 \eta} (1 + \lambda_1)^{d/2} \lambda_1}{(2\pi)^{\frac{d}{2}}} \times \left[ \int_{z \in R^{d(d+1)/2}} e^{-\frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} z|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} z - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right]} dz \right]^{-1}. \end{aligned} \quad (15)$$

The constants  $H_\lambda$  and  $H_{\lambda_1}$  ensure that  $h_{0,t}$  and  $h_{1,t}$  are properly normalized densities.

**Understanding the measure  $Q$ .** The measure  $Q$  is designed such that the distribution of  $f$  under the measure  $Q$  is approximately the conditional distribution of  $f$  given  $\mathcal{I}(T) > b$ . The two



terms corresponding to the probabilities  $\rho_1$  and  $\rho_2$  are included to ensure the absolute continuity and to control the tail of the likelihood ratio. Thus,  $\rho_1$  and  $\rho_2$  will be sent to zero eventually.

We now provide an explanation of the leading term corresponding to the probability  $1 - \rho_1 - \rho_2$ . To understand  $h_{0,t}$ , we use the notation  $\alpha_t$  in (8) and rewrite the density function as

$$h_{0,t}(f(t), \partial f(t), \partial^2 f(t)) \propto \mathbb{I}_{A_t} \exp\{-\lambda u_t(\alpha_t - u_t)\} \times \exp\left\{-\frac{1-\lambda}{2}|\partial f(t)|^2\right\} \\ \times \exp\left\{-\frac{1}{2}\left[\frac{|\mu_{20}\mu_{22}^{-1}\bar{f}_t''|^2}{1-\mu_{20}\mu_{22}^{-1}\mu_{02}} + \left|\mu_{22}^{-1/2}\bar{f}_t'' - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma}\right|^2\right]\right\},$$

which factorizes into three pieces consisting of  $\alpha_t$ ,  $\partial f(t)$ , and  $\bar{f}_t''$  respectively. We consider the change of variables from  $(f(t), \partial f(t), \partial^2 f(t))$  to  $(\alpha_t, \partial f(t), \bar{f}_t'')$ . Then, under the distribution  $h_{0,t}$ , the random vectors  $\alpha_t$ ,  $\partial f(t)$ , and  $\bar{f}_t''$  are independent. Note that  $h_{0,t}$  is defined on the set  $A_t = \{\alpha_t > u_t - \eta u_t^{-1}\}$  where  $\eta$  will be sent to zero eventually. Then,  $\alpha_t - u_t$  is approximately an exponential random variable with rate  $\lambda u_t$ ;  $\partial f(t)$  and  $\bar{f}_t''$  are two independent Gaussian random vectors. The density  $h_{1,t}$  has a similar interpretation. The only difference is that  $h_{1,t}$  is defined on the set  $\{\alpha_t - u_t < -\eta u_t^{-1}\}$  and  $u_t - \alpha_t$  follows approximately an exponential distribution. For the last piece corresponding to  $\rho_2$ , the density is simply an exponential tilting of  $f(t)$ .

Under the dominating mixture component, to generate an  $f(t)$  from  $Q$ , a random index  $\tau$  is first sampled from  $T$  following density  $l(t)$ , then  $(f(\tau), \partial f(\tau), \partial^2 f(\tau))$  is sampled according to  $h_{0,\tau}$ . This implies that the large value of the integral  $\int_T e^{\mu(t)+\sigma f(t)} dt$  is mostly caused by the fact that the field reaches a high level at  $\tau$ ; more precisely,  $\alpha_\tau$  reaches a high level of  $u_\tau$  (with an exponential overshoot of rate  $\lambda u_\tau$ ). Therefore, the random index  $\tau$  localizes the position where the field  $\alpha_t$  goes very high. The distribution of  $\tau$  given as in (14) is very concentrated around  $t_*$ . This suggests that the maximum of  $\alpha_t$  (or  $f(t)$ ) is attained within  $O_p(u^{-1/2})$  distance from  $t_*$ .

We now consider the case where  $\mu(t) \equiv 0$ . We choose  $l(t)$  to be the uniform distribution over set  $T$  and have that

$$\frac{dQ}{dP} = (1 - \rho_1 - \rho_2) \int_T \frac{1}{\text{mes}(T)} \cdot LR(t) dt + \rho_1 \int_T \frac{1}{\text{mes}(T)} \cdot LR_1(t) dt + \rho_2 \int_T \frac{1}{\text{mes}(T)} \cdot LR_2(t) dt, \quad (16)$$

where  $\text{mes}(\cdot)$  is the Lebesgue measure. The following theorem states that  $Q$  is a good approximation of  $P_b^*$  with appropriate choice of the tuning parameters.

**Theorem 3** *Consider a Gaussian random field  $\{f(t) : t \in T\}$  living on a domain  $T$  satisfying conditions C1-6. Then, for any  $\varepsilon > 0$ , there exist  $b_0, \eta, \rho_1, \rho_2, \lambda$  and  $\lambda_1$  such that  $Q$  approximates  $P_b^*$  in total variation, that is, for all  $b > b_0$*

$$\sup_{A \in \mathcal{F}} |Q(A) - P_b^*(A)| \leq \varepsilon.$$

**Remark 4** *In particular, as  $\varepsilon$  becomes small, we will send  $\eta, \rho_1, \rho_2 \rightarrow 0$ ,  $\lambda \rightarrow 1-$ , and  $\lambda_1$  stays as a constant. Theorem 3 is the central result of this paper. We present its detailed proof. The technical developments of other theorems are all based on that of Theorem 3. Therefore, we only layout their key steps and the major differences from that of Theorem 3.*

**Remark 5** *The measure corresponding to the last mixture component in (11),  $\int_T \frac{LR_2(t)}{\text{mes}(T)} dt$ , has been employed by [36] to develop approximations for  $v(b)$ . We emphasize that the measure constructed in this paper is substantially different. In fact, the measure corresponding to  $LR_2(t)$  does not appear in the main proof. We included it to control the tail of the likelihood ratio in one lemma.*



To illustrate the application of the measure  $Q$ , we provide a further characterization of the conditional distribution  $P_b^*$  by presenting another approximation result which is easier to understand at an intuitive level. Let

$$\gamma_u(t) \triangleq f(t) + \frac{\mathbf{1}^\top \tilde{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} + \mu_\sigma(t), \quad \beta_u(T) = \sup_{t \in T} \gamma_u(t), \quad \tilde{P}_b(f(\cdot) \in A) = P(f(\cdot) \in A | \beta_u(T) > u).$$

The process  $\gamma_u(t)$  is slightly different from  $\alpha_t$ . The following theorem states that the measure  $Q$  also approximates the distribution  $\tilde{P}_b$  in total variation for  $b$  large.

**Theorem 6** *Consider a Gaussian random field  $\{f(t) : t \in T\}$  living on a domain  $T$  satisfying conditions C1-6. Then, for any  $\varepsilon$ , with the same choice of tuning parameters as in Theorem 3,  $Q$  approximates  $\tilde{P}_b$  in total variation, that is, for all  $b > b_0$*

$$\sup_{A \in \mathcal{F}} |Q(A) - \tilde{P}_b(A)| \leq \varepsilon.$$

### 2.3 Some implications of the theorems

The results of Theorems 3 and 4 provide both qualitative and quantitative descriptions of  $P_b^*$ . From a qualitative point of view, Theorems 3 and 4 suggest that

$$\sup_{A \in \mathcal{F}} |P_b^*(A) - \tilde{P}_b(A)| \rightarrow 0, \tag{17}$$

as  $b \rightarrow \infty$ . Note that  $\gamma_u(t)$  itself is a Gaussian process. Thus, the above convergence results connect the tail events of the exponential integral to the supremum of another Gaussian process that is a linear combination of  $f$  and its derivative field. We emphasize that classic results (e.g. those in the introduction) do not (directly) apply to  $\gamma_u(t)$  in that  $\gamma_u(t)$  is indexed by  $u$  (a function of  $b$ ). Nonetheless, the result in (17) together with our previous explanation of the measure  $Q$  provides a qualitative description of  $P_b^*$  given that the high excursion of the supremum of Gaussian processes is much more intensively studied in the literature than the integral  $\mathcal{I}(T)$  and the distribution of  $f$  under  $Q$  has a closed form representation.

From a quantitative point of view, Theorem 3 implies that for any bounded function  $\Xi : C(T) \rightarrow \mathbb{R}$  the conditional expectation  $E[\Xi(f) | \mathcal{I}(T) > b]$  can be approximated by  $E^Q[\Xi(f)]$ , more precisely,

$$E[\Xi(f) | \mathcal{I}(T) > b] - E^Q[\Xi(f)] \rightarrow 0 \tag{18}$$

as  $b \rightarrow \infty$ . The expectation  $E^Q[\Xi(f)]$  is much easier to compute (both analytically and numerically) via the following identity

$$E^Q[\Xi(f)] = E^Q [ E[\Xi(f) | \imath, \tau, f(\tau), \partial f(\tau), \partial^2 f(\tau)] ]. \tag{19}$$

Note that the inner expectation is under the measure  $P$  in that the conditional distribution of  $f$  given  $(f(\tau), \partial f(\tau), \partial^2 f(\tau))$  under  $Q$  is the same as that under  $P$ . Furthermore, conditional on  $(f(\tau), \partial f(\tau), \partial^2 f(\tau))$ , the process  $f(t)$  is also a Gaussian process and has the expansion

$$f(t) = f(\tau) + \partial f(\tau)^\top (t - \tau) + \frac{1}{2} (t - \tau)^\top \Delta f(\tau) (t - \tau) + o(|t - \tau|^2).$$

These results provide sufficient tools to evaluate the conditional expectation

$$E [\Xi(f) | \imath, \tau, f(\tau), \partial f(\tau), \partial^2 f(\tau)].$$

Once the above expectation has been evaluated, we may proceed to the outer expectation in (19). Note that the inner expectation is a function of  $(\iota, \tau, f(\tau), \partial f(\tau), \partial^2 f(\tau))$ , the joint distribution of which is in a closed form. Thus, evaluating the outer expectation is usually an easier task. In fact, the proof of Theorem 3 is an exercise of the above strategy by considering that  $\Xi(f) = (dP/dQ)^2$ .

**Remark 7** *According to the detailed proof of Theorem 3, the approximation (18) is applicable to all the functions such that  $\sup_b E[\Xi^2(f)|\mathcal{I}(T) > b] < \infty$ . To see that, we need to change the statement and the proof of Lemma 12 presented in Section 3.*

## 2.4 Efficient Rare-event Simulation for $\mathcal{I}(T)$

In the preceding subsection we constructed a change of measure that asymptotically approximates the conditional distribution of  $f$  given  $\mathcal{I}(T) > b$ . In this section, we construct an efficient importance sampling estimator based on this change of measure to compute  $v(b)$  as  $b \rightarrow \infty$ . We evaluate the overall computation efficiency using a concept that has its root in the general theory of computation in both continuous and discrete settings [39, 46]. In particular, completely analogous notions in the setting of complexity theory of continuous problems lead to the notion of tractability of a computational problem [48].

**Definition 8** *A Monte Carlo estimator is said to be a fully polynomial randomized approximation scheme (FPRAS) for estimating  $v(b)$  if, for some  $q_1, q_2$ , and  $d > 0$ , it outputs an averaged estimator that is guaranteed to have at most  $\varepsilon > 0$  relative error with confidence at least  $1 - \delta \in (0, 1)$  in  $O(\varepsilon^{-q_1} \delta^{-q_2} |\log v(b)|^d)$  function evaluations.*

Equivalently, one needs to compute an estimator  $Z_b$  with complexity  $O(\varepsilon^{-q_1} \delta^{-q_2} |\log v(b)|^d)$  such that

$$P(|Z_b/v(b) - 1| > \varepsilon) < \delta. \quad (20)$$

In the literature of rare-event simulations, an estimator  $L_b$  is said to be *strongly efficient* in estimating  $v(b)$  if  $EL_b = v(b)$  and  $\sup_b \text{Var} L_b / v^2(b) < \infty$ . Suppose that a strongly efficient estimator  $L_b$  has been obtained. Let  $\{L_b^{(j)} : j = 1, \dots, n\}$  be i.i.d. copies of  $L_b$ . The averaged estimator

$$Z_b = \frac{1}{n} \sum_{j=1}^n L_b^{(j)}$$

has a *relative* mean squared error equal to  $\sqrt{E(Z_b/v(b) - 1)^2} = \text{Var}^{1/2}(L_b) n^{-1/2} v(b)^{-1}$ . A simple consequence of Chebyshev's inequality yields

$$P(|Z_b/v(b) - 1| \geq \varepsilon) \leq \frac{\text{Var}(L_b)}{\varepsilon^2 n v^2(b)}.$$

Thus, it suffices to simulate  $n = O(\varepsilon^{-2} \delta^{-1})$  i.i.d. replicates of  $L_b$  to achieve the accuracy in (20).

The so-called importance sampling is based on the identity  $P(A) = E^Q[\mathbb{I}_A dP/dQ]$ . The random variable  $\mathbb{I}_A dP/dQ$  is an unbiased estimator of  $P(A)$ . It is well known that if one chooses  $Q(\cdot) = P(\cdot|A)$  then  $\mathbb{I}_A dP/dQ$  has zero variance. The measure  $Q$  created in the previous subsection is a good approximation of  $P_b^*$  and thus it naturally leads an estimator for  $v(b)$  with small variance.

In addition to the variance control, another issue is that the random fields considered in this paper are continuous objects. Computer can only perform discrete simulations. Thus, we must use a discrete object approximating the continuous field to implement the algorithms. The bias

caused by the discretization must be well controlled relative to  $v(b)$ . In addition, the complexity of generating one such discrete object should also be considered in order to control the overall computational complexity to achieve an FPRAS.

To start with, we create a regular lattice covering  $T$  in the following way. Let  $G_{N,d}$  be a subset of  $R^d$  containing

$$G_{N,d} = \left\{ \left( \frac{i_1}{N}, \frac{i_2}{N}, \dots, \frac{i_d}{N} \right) : i_1, \dots, i_d \in \mathbb{Z} \right\}.$$

That is,  $G_{N,d}$  is a regular lattice on  $R^d$ . For each  $t = (t^1, \dots, t^d) \in G_{N,d}$ , define

$$T_N(t) = \left\{ (s^1, \dots, s^d) \in T : s^j \in (t^j - 1/N, t^j] \text{ for } j = 1, \dots, d \right\}$$

that is the  $\frac{1}{N}$ -cube intersected with  $T$  and cornered at  $t$ . Furthermore, let

$$T_N = \{t \in G_{N,d} : T_N(t) \neq \emptyset\}. \quad (21)$$

Since  $T$  is compact,  $T_N$  is a finite set. We enumerate the elements in  $T_N = \{t_1, \dots, t_M\}$ , where  $M = O(N^d)$ . We further define

$$X = (X_1, \dots, X_M)^\top \triangleq (f(t_1), \dots, f(t_M))^\top$$

and use

$$v_M(b) = P(\mathcal{I}_M(T) > b)$$

as an approximation of  $v(b)$  where

$$\mathcal{I}_M(T) = \sum_{i=1}^M \text{mes}(T_N(t_i)) \times e^{\sigma X_i + \mu(t_i)}. \quad (22)$$

We have the following theorem to control the bias.

**Theorem 9** *Consider a Gaussian random field  $f$  satisfying conditions in Theorem 3. For any  $\varepsilon_0 > 0$ , there exists  $\kappa_0$  such that for any  $\varepsilon \in (0, 1)$ , if  $N \geq \kappa_0 \varepsilon^{-1-\varepsilon_0} (\log b)^{2+\varepsilon_0}$ , then for  $b > 2$*

$$\frac{|v_M(b) - v(b)|}{v(b)} < \varepsilon.$$

We estimate  $v_M(b)$  using a discrete version of the change of measure proposed in the previous section. The specific algorithm is given as follows.

1. Generate a random indicator  $\iota \sim \text{Bernoulli}(\rho_2)$ . If  $\iota = 1$ , then
  - (a) Generate  $\iota$  uniformly from  $\{1, \dots, M\}$ .
  - (b) Generate  $X_\iota \sim N(u_\iota, 1)$ .
  - (c) Given  $(t_\iota, X_\iota)$ , simulate the joint field  $(f(t), \partial f(t), \partial^2 f(t))$  on the lattice  $T_N \setminus \{t_\iota\}$  from the original conditional distribution under  $P$ .
2. If  $\iota = 0$ 
  - (a) If  $\mu(t)$  is not constant, simulate a random index  $\iota$  proportional to  $l(t_\iota)$ , that is,  $P(\iota = i) = l(t_i)/\kappa$  and  $\kappa = \sum_{i=1}^M l(t_i)$ ; if  $\mu(t) \equiv 0$ , then  $\iota$  is simulated uniformly over  $\{1, \dots, M\}$ .

- (b) Given the realized  $\iota$ , simulate  $f(t_\iota) = X_\iota = x, \partial f(t_\iota) = y, \partial^2 f(t_\iota) = z$  from density function

$$h_{all}(x, y, z) = \frac{1 - \rho_1 - \rho_2}{1 - \rho_2} h_{0, t_\iota}(x, y, z) + \frac{\rho_1}{1 - \rho_2} h_{1, t_\iota}(x, y, z).$$

- (c) Given  $(t_\iota, f(t_\iota), \partial f(t_\iota), \partial^2 f(t_\iota))$ , simulate the joint field  $(f(t), \partial f(t), \partial^2 f(t))$  on the lattice  $T_N \setminus \{t_\iota\}$  from the original conditional distribution under  $P$ .

### 3. Output

$$\tilde{L}_b = \frac{\mathbb{I}_{\{\mathcal{I}_M(T) > b\}}}{\frac{1 - \rho_1 - \rho_2}{\kappa} \sum_{i=1}^M l(t_i) LR(t_i) + \frac{\rho_1}{\kappa} \sum_{i=1}^M l(t_i) LR_1(t_i) + \rho_2 \sum_{i=1}^M \frac{1}{M} LR_2(t_i)}. \quad (23)$$

Let  $Q_M$  be the measure induced by the above simulation scheme. Then, it is not hard to verify that  $\tilde{L}_b = \mathbb{I}_{\{\mathcal{I}_M(T) > b\}} dP/dQ_M$  and thus  $\tilde{L}_b$  is an unbiased estimator of  $v_M(b)$ . The next theorem states the strong efficiency of the above algorithm.

**Theorem 10** *Suppose  $f$  is a Gaussian random field satisfying conditions in Theorem 3. If  $N$  is chosen as in Theorem 9 and all the other parameters are chosen as in Theorem 3, then there exists some constant  $\kappa_1 > 0$  such that*

$$\sup_{b > 1} \frac{E^{Q_M} \tilde{L}_b^2}{v_M^2(b)} \leq \kappa_1.$$

Let  $Z_b$  be the average of  $n$  i.i.d. copies of  $\tilde{L}_b$ . According to the results in Theorem 9, we have that

$$\left| \frac{Z_b}{v(b)} - 1 \right| \leq \left| \frac{Z_b}{v_M(b)} (v_M(b)/v(b) - 1) \right| + \left| \frac{Z_b}{v_M(b)} - 1 \right| \leq \varepsilon \left| \frac{Z_b}{v_M(b)} \right| + \left| \frac{Z_b}{v_M(b)} - 1 \right|.$$

The results of Theorem 10 suggest that

$$P(|Z_b/v_M(b) - 1| \geq \varepsilon) \leq \frac{\kappa_1}{\varepsilon^2 n}.$$

If we choose  $n = \kappa_1 \varepsilon^{-2} \delta^{-1}$ , then

$$P(|Z_b/v(b) - 1| \geq 3\varepsilon) \leq \delta.$$

Thus, the accuracy level as in (20) has been achieved. Note that simulating one  $\tilde{L}_b$  consists of generating a multivariate Gaussian random vector of dimension  $M \times (d+1)(d+2)/2 = O(N^d) = O((\log b)^{(2+\varepsilon_0)d} \varepsilon^{-(1+\varepsilon_0)d})$ . The complexity of generating such a vector is at the most  $O(N^3)$ . Thus, the overall complexity is  $O(\varepsilon^{-2-(3+3\varepsilon_0)d} \delta^{-1} (\log b)^{(6+3\varepsilon_0)d})$ . The proposed estimator in (23) is a FPRAS.

**Remark 11** *The proposed algorithm can also be used to compute conditional expectations via the representation  $E[\Xi(f)|\mathcal{I}(T) > b] = E[\Xi(f); \mathcal{I}(T) > b]/v(b)$ , where  $E[\Xi(f); \mathcal{I}(T) > b]$  can be estimated by  $\Xi(f)dP/dQ_M$  and  $v(b)$  can be estimated by  $\mathbb{I}_{\{\mathcal{I}(T) > b\}} dP/dQ_M$ .*

### 3 Proof of Theorem 3

We use the following simple yet powerful lemma to prove Theorem 3.

**Lemma 12** *Let  $Q_0$  and  $Q_1$  be probability measures defined on the same  $\sigma$ -field  $\mathcal{F}$  such that  $dQ_1 = r^{-1}dQ_0$  for a positive random variable  $r$ . Suppose that for some  $\varepsilon > 0$ ,  $E^{Q_1}[r^2] = E^{Q_0}[r] \leq 1 + \varepsilon$ . Then,*

$$\sup_{|X| \leq 1} |E^{Q_1}(X) - E^{Q_0}(X)| \leq \varepsilon^{1/2}.$$

**Proof of Lemma 12.**

$$\begin{aligned} |E^{Q_1}(X) - E^{Q_0}(X)| &= |E^{Q_1}[(1-r)X]| \\ &\leq E^{Q_1}|r-1| \leq [E^{Q_1}(r-1)^2]^{1/2} = (E^{Q_1}[r^2] - 1)^{1/2} \leq \varepsilon^{1/2}. \end{aligned}$$

■

We also need the following approximations for the tail probability  $v(b)$  (see Theorem 3.4 and Corollary 3.5 in [36]).

**Proposition 13** *Consider a Gaussian random field  $\{f(t) : t \in T\}$  living on a domain  $T$  satisfying conditions C1-6. If  $\mu(t)$  has one unique maximum in  $T$  denoted by  $t_*$ , then*

$$v(b) \sim (2\pi)^{d/2} \det(\Delta\mu_\sigma(t_*))^{-1/2} G(t_*) \cdot u^{d/2-1} \exp\left\{-\frac{(u - \mu_\sigma(t_*))^2}{2}\right\},$$

where  $u$  is as defined in (5),  $G(t)$  is defined as

$$\frac{\det(\Gamma)^{-\frac{1}{2}}}{(2\pi)^{\frac{(d+1)(d+2)}{4}}} e^{\frac{\mathbf{1}^T \mu_{22}^{-1} \mathbf{1}}{8\sigma^2} + Bt} \cdot \int_{z \in R^{d(d+1)/2}} \exp\left\{-\frac{1}{2} \left[ \frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left| \mu_{22}^{-1/2}z - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right]\right\} dz.$$

If  $\mu(t) \equiv 0$ ,  $G(t)$  is a constant denoted by  $G$ . Then,

$$v(b) \sim \text{mes}(T)G \cdot u^{d-1} \exp\left\{-\frac{u^2}{2}\right\}.$$

#### 3.1 Case 1: $\mu(t)$ is not a constant

To make the proof smooth, we arrange the statement of the rest supporting lemmas in Section 7. We start the proof of Theorem 3 when  $\mu(t)$  is not a constant. Note that

$$E^Q \left[ \left( \frac{dP_b^*}{dQ} \right)^2 \right] = v(b)^{-2} E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b \right].$$

Thanks to Lemma 12, we only need to show that as  $b \rightarrow \infty$

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b \right] = E^Q \left[ E_{i,\tau}^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b \right] \right] \leq (1 + \varepsilon) v(b)^2.$$

where we use the notation  $E_{i,\tau}^Q[\cdot] = E^Q[\cdot \mid i, \tau]$  to denote the conditional expectation given  $i$  and  $\tau$ .  $\tau \in T$  is the random index described as in the simulation scheme admitting a density function  $l(t)$  if  $i = 0$  and  $mes^{-1}(T)\mathbb{I}_T(t)$  if  $i = 1$ . Note that

$$E_{i,\tau}^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b \right] = E_{i,\tau}^Q \left[ E_{i,\tau}^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b \mid f(\tau), \partial f(\tau), \partial^2 f(\tau) \right] \right].$$

For the rest of the proof, we mostly focus on the conditional expectation

$$E_{i,\tau}^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b \mid f(\tau), \partial f(\tau), \partial^2 f(\tau) \right].$$

The rest of the discussion is conditional on  $i$  and  $\tau$ . To simplify the notations, for a given  $\tau$ , we define

$$f_*(t) = f(t) - u_\tau C(t - \tau).$$

On the set  $\{\mathcal{I}(T) > b\}$ ,  $f(\tau)$  reaches a level  $u_\tau$ , and  $E[f(t) \mid f(\tau) = u_\tau] = u_\tau C(t - \tau)$ . Thus,  $f_*(t)$  is the field with the conditional expectation removed. From now on, we work with this shifted field  $f_*(t)$ . Correspondingly, we have

$$\partial f_*(t) = \partial f(t) - u_\tau \partial C(t - \tau), \quad \partial^2 f_*(t) = \partial^2 f(t) - u_\tau \partial^2 C(t - \tau).$$

We further define notations

$$\begin{aligned} w &= f_*(\tau), \quad y = \partial f_*(\tau), \quad z = \partial^2 f_*(\tau), \quad \mathbf{z} = \Delta f_*(\tau), \\ \tilde{y} &= \partial f_*(\tau) + \partial \mu_\sigma(\tau), \quad \tilde{\mathbf{z}} = \Delta f_*(\tau) + \mu_\sigma(\tau)I + \Delta \mu_\sigma(\tau), \\ w_t &= f_*(t), \quad y_t = \partial f_*(t), \quad z_t = \partial^2 f_*(t), \quad \bar{z}_t = \partial^2 f_*(t) - u_t \mu_{02}. \end{aligned} \quad (24)$$

Under the measure  $Q$  and a given  $\tau$ , if  $i = 0$ ,  $(w, y, z)$  has density function

$$h_{all}^*(w, y, z) = \frac{1 - \rho_1 - \rho_2}{1 - \rho_2} h_{0,\tau}^*(w, y, z) + \frac{\rho_1}{1 - \rho_2} h_{1,\tau}^*(w, y, z); \quad (25)$$

if  $i = 1$ , then  $(w, y, z)$  follows density  $h_\tau^*(w, y, z)$ . The forms of the densities can be derived from  $h_{0,t}$ ,  $h_{1,t}$ , and  $h$ . In particular, their expressions are given as follows

$$\begin{aligned} h_{0,\tau}^*(w, y, z) &\propto \mathbb{I}_{A_\tau} \times \exp \left\{ -\lambda u_\tau \left( w + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} + \frac{B_\tau}{u_\tau} \right) - \frac{1}{2} |y|^2 \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} z|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} z - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right] \right\} \\ h_{1,\tau}^*(w, y, z) &\propto \mathbb{I}_{A_\tau^c} \times \exp \left\{ \lambda_1 u_\tau \left( w + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} + \frac{B_\tau}{u_\tau} \right) - \frac{1}{2} |y|^2 \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} z|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} z - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right] \right\} \\ h_\tau^*(w, y, z) &= h(w, y, z) = \frac{\det(\Gamma)^{-\frac{1}{2}}}{(2\pi)^{\frac{(d+1)(d+2)}{4}}} e^{-\frac{1}{2} \left[ y^\top y + \frac{|w - \mu_{20} \mu_{22}^{-1} z|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + z^\top \mu_{22}^{-1} z \right]}, \end{aligned}$$



and  $A_\tau = \left\{ w + \frac{y^\top y}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} + \frac{B_\tau}{u_\tau} > -\eta u_\tau^{-1} \right\}$  is defined as in (7).

In the next step, we will compute  $dQ/dP$  in the form of  $f_*(t)$ . Basically, we replace  $f(t)$  by  $f_*(t) + u_\tau C(t - \tau)$ ,  $\partial f(t)$  by  $y_t + u_\tau \partial C(t - \tau)$ ,  $\partial^2 f(t)$  by  $z_t + u_\tau \partial^2 C(t - \tau)$ , and  $\bar{f}_t'' = \partial^2 f(t) - u_t \mu_{02}$  by  $\bar{z}_t + u_\tau \partial^2 C(t - \tau)$ . For the likelihood ratio terms  $LR$  and  $LR_1$  in (12), note that the  $|\partial f(t)|^2$  terms in  $h_{0,t}$  and  $h_{1,t}$  cancel with those in  $h(f(t), \partial f(t), \partial^2 f(t))$ , that is,

$$LR(t) = \mathbb{I}_{A_t} \cdot \frac{H_\lambda \cdot u_t e^{-\lambda u_t \left( f(t) + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} - u_t \right) - \frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} \bar{f}_t''|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} \bar{f}_t'' - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right]}}{\frac{\det(\Gamma)^{-\frac{1}{2}}}{(2\pi)^{\frac{(d+1)(d+2)}{4}}} e^{-\frac{1}{2} \left[ \frac{(f(t) - \mu_{20} \mu_{22}^{-1} \partial^2 f(t))^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \partial^2 f(t)^\top \mu_{22}^{-1} \partial^2 f(t) \right]}}$$

We insert the notations in (24) and obtain that

$$\begin{aligned} LR(t) &= \mathbb{I}_{A_t} \cdot u_t H_\lambda \exp \left\{ -\lambda u_t \left( w_t + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top (\bar{z}_t + \mu_2(t - \tau) u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} - u_t \right) \right\} \\ &\times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} (\bar{z}_t + \mu_2(t - \tau) u_\tau)|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} (\bar{z}_t + \mu_2(t - \tau) u_\tau) - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right] \right\} \\ &\times h_{x,z}^{-1} (w_t + u_\tau C(t - \tau), z_t + u_\tau \partial^2 C(t - \tau)), \end{aligned} \quad (26)$$

where

$$h_{x,z}(x, z) = \frac{\det(\Gamma)^{-\frac{1}{2}}}{(2\pi)^{\frac{(d+1)(d+2)}{4}}} e^{-\frac{1}{2} \left[ \frac{(x - \mu_{20} \mu_{22}^{-1} z)^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + z^\top \mu_{22}^{-1} z \right]}, \quad (27)$$

which is the function  $h(x, y, z)$  with the  $|y|^2$  term removed. Similarly, we have that

$$\begin{aligned} LR_1(t) &= \mathbb{I}_{A_t^c} \cdot u_t H_{\lambda_1} \exp \left\{ \lambda_1 u_t \left( w_t + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top (\bar{z}_t + \mu_2(t - \tau) u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} - u_t \right) \right\} \\ &\times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20} \mu_{22}^{-1} (\bar{z}_t + \mu_2(t - \tau) u_\tau)|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + \left| \mu_{22}^{-1/2} (\bar{z}_t + \mu_2(t - \tau) u_\tau) - \frac{\mu_{22}^{1/2} \mathbf{1}}{2\sigma} \right|^2 \right] \right\} \\ &\times h_{x,z}^{-1} (w_t + u_\tau C(t - \tau), z_t + u_\tau \partial^2 C(t - \tau)). \end{aligned} \quad (28)$$

With the analytic forms (26) and (28), we proceed to the likelihood ratio in (11)

$$\frac{dQ}{dP} = (1 - \rho_1 - \rho_2)K + \rho_1 K_1 + \rho_2 K_2 \quad (29)$$

where

$$K = \int_{A^*} l(t) LR(t) dt, \quad K_1 = \int_{(A^*)^c} l(t) LR_1(t) dt, \quad K_2 = \frac{1}{\text{mes}(T)} \int_T e^{-\frac{1}{2} u_t^2 + u_t w_t + u_t u_\tau C(t - \tau)} dt.$$

where the set  $A^*$  (depending on the sample path  $f_*(t)$ ) is defined as

$$A^* = \left\{ t : w_t + C(t - \tau) u_\tau + \frac{|y_t + u_\tau \cdot \partial C(t - \tau)|^2}{2u_t} + \frac{\mathbf{1}^\top (\bar{z}_t + u_\tau \mu_2(t - \tau))}{2\sigma u_t} + \frac{B_t}{u_t} > u_t - \frac{\eta}{u_t} \right\}. \quad (30)$$

We may equivalently define  $A^* = \{t : f \in A_t\}$ . Note that  $LR(t) = 0$  if  $f \notin A_t$ . Thus, the integral of  $K$  is essentially on the set  $A^*$  and  $K_1$  is on the compliment of  $A^*$ .

Furthermore, we have that

$$\begin{aligned} E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b \right] &\leq E^Q \left\{ E_{i,\tau}^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2} ; \mathcal{I}(T) > b \right] \right\} \\ &\leq E^Q \left\{ E_{i,\tau}^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K]^2} ; \mathcal{I}(T) > b, \mathcal{A}_\tau \geq 0 \right] \right\} \\ &\quad + E^Q \left\{ E_{i,\tau}^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2} ; \mathcal{I}(T) > b, \mathcal{A}_\tau < 0 \right] \right\}, \end{aligned} \quad (31)$$

where

$$\mathcal{A}_\tau = w + \frac{y^\top y}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} + \frac{B_\tau}{u_\tau}. \quad (32)$$

Note that the term  $K_2$  is not used in the main analysis. In fact,  $K_2$  is only used in Lemma 16 for the purpose of localization that will be presented later. The rest of the analysis consists of three main parts.

**Part 1** Conditional on  $(i, \tau, f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau))$ , we study the event

$$\mathcal{E}_b = \{\mathcal{I}(T) > b\}, \quad (33)$$

and write the occurrence of this event almost as a deterministic function of  $f_*(\tau)$ ,  $\partial f_*(\tau)$ , and  $\partial^2 f_*(\tau)$ , equivalently  $(w, y, z)$ .

**Part 2** Conditional on  $(i, \tau, f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau))$ , we express  $K$  and  $K_1$  as functions of  $f_*(\tau)$ ,  $\partial f_*(\tau)$ ,  $\partial^2 f_*(\tau)$  with small correction terms.

**Part 3** We combine the results from the first two parts and obtain an approximation of (31).

All the subsequent derivations are conditional on  $i$  and  $\tau$ .

### 3.1.1 Preliminary calculations

To proceed, we provide the Taylor expansions for  $f_*(t)$ ,  $C(t)$ , and  $\mu(t)$ .

- Expansion of  $f_*(t)$  given  $(f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau))$ . Let  $t - \tau = ((t - \tau)_1, \dots, (t - \tau)_d)$ . Conditional on  $(f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau))$ , we first expand the random function

$$\begin{aligned} f_*(t) &= E[f_*(t) | f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau)] + g(t - \tau) \\ &= f_*(\tau) + \partial f_*(\tau)^\top (t - \tau) + \frac{1}{2} (t - \tau)^\top \Delta f_*(\tau) (t - \tau) \\ &\quad + g_3(t - \tau) + R_f(t - \tau) + g(t - \tau), \end{aligned} \quad (34)$$

where

$$g_3(t - \tau) = \frac{1}{6} \sum_{i,j,k} E[\partial_{ijk}^3 f_*(\tau) | f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau)] (t - \tau)_i (t - \tau)_j (t - \tau)_k,$$

is the third order expansion. Note that  $\partial_{ijk}^3 f_*(\tau)$  is independent of  $(f_*(\tau), \Delta f_*(\tau))$  and

$$E [\partial_{ijk}^3 f_*(\tau) | f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau)] = - \sum_l \partial_{ijkl}^4 C(0) \partial_l f_*(\tau).$$

$g(t)$  is a mean zero Gaussian random field such that  $Eg^2(t) = O(|t|^6)$  as  $t \rightarrow 0$ . In addition, the distribution of  $g(t)$  is independent of  $\tau, f_*(\tau), \partial f_*(\tau)$ , and  $\partial^2 f_*(\tau)$ .  $R_f(t - \tau) = O(|t - \tau|^4)$  is the remainder term of the Taylor expansion of  $E[f_*(t) | f_*(\tau), \partial f_*(\tau), \partial^2 f_*(\tau)]$ .

- Expansion of  $C(t)$ :

$$C(t) = 1 - \frac{1}{2} t^\top t + C_4(t) + R_C(t), \quad (35)$$

where  $R_C(t) = O(|t|^6)$  and  $C_4(t) = \frac{1}{24} \sum_{ijkl} \partial_{ijkl}^4 C(0) t_i t_j t_k t_l$ .

- Expansion of  $\mu(t)$ :

$$\mu(t) = \mu(\tau) + \partial \mu(\tau)^\top (t - \tau) + \frac{1}{2} (t - \tau)^\top \Delta \mu(\tau) (t - \tau) + R_\mu(t - \tau), \quad (36)$$

where  $R_\mu(t - \tau) = O(|t - \tau|^3)$ .

We write

$$R(t) = R_f(t) + u_\tau R_C(t) + R_\mu(t)/\sigma$$

to denote all the remainder terms. For  $0 < \epsilon < \delta$  sufficiently small, let

$$\begin{aligned} \mathcal{L} = & \left\{ |\tau - t_*| < u^{-1/2+\epsilon}, |f(\tau) - u_\tau| \leq u^{1/2+\epsilon}, |\partial f(\tau)| < u^\epsilon, |\partial^2 f(\tau) - u_\tau \mu_{02}| < u^\epsilon, \right. \\ & \left. \sup_{|t-\tau| < u^{-1/2+\delta}} |\partial^3 f(t) - u_\tau \partial^3 C(t - \tau)| < u^{1/2+\epsilon}, \sup_{|t-\tau| < u^{-1/2-\delta}} |g(t)| < u^{-1-\delta} \right\}. \end{aligned} \quad (37)$$

By Lemma 16 whose proof uses the last component  $LR_2(t)$ , we have that

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{E}_b, \mathcal{L}^c \right] = o(1) v^2(b).$$

Therefore we only need to consider the second moment on the set  $\mathcal{L}$ , that is,

$$\begin{aligned} E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{E}_b, \mathcal{L} \right] & \leq E^Q \left[ E_{i,\tau}^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau > 0 \right] \right] \\ & + E^Q \left[ E_{i,\tau}^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0 \right] \right], \end{aligned} \quad (38)$$

where  $K$  and  $K_1$  are given as in (29). Note that we can rewrite the localization set in terms of  $f_*(t)$  as

$$\begin{aligned} \mathcal{L} = & \left\{ |\tau - t_*| < u^{-1/2+\epsilon}, |w| \leq u^{1/2+\epsilon}, |y| < u^\epsilon, |z| < u^\epsilon, \right. \\ & \left. \sup_{|t-\tau| < u^{-1/2+\delta}} |\partial^3 f_*(t)| < u^{1/2+\epsilon}, \sup_{|t-\tau| < u^{-1/2-\delta}} |g(t)| < u^{-1-\delta} \right\}. \end{aligned} \quad (39)$$

We will focus on the terms on the right-hand-side of (38) in the subsequent derivations. Now, we start to carry out each part of the program.

### 3.2 Part 1

All the derivations in this part are conditional on specific values of  $\iota$ ,  $\tau$ ,  $f_*(\tau)$ ,  $\partial f_*(\tau)$ , and  $\partial^2 f_*(\tau)$ , equivalently,  $\iota$ ,  $\tau$ ,  $w$ ,  $y$ , and  $z$ . By definition

$$\mathcal{I}(T) = \int_T e^{\sigma f_*(t) + \sigma u_\tau C(t-\tau) + \mu(t)} dt.$$

We insert the expansions in (34), (35), and (36) into the expression of  $\mathcal{I}(T)$  and obtain that

$$\begin{aligned} \mathcal{I}(T) &= \int_{t \in T} \exp \left\{ \sigma \left[ w + y^\top (t - \tau) + \frac{1}{2} (t - \tau)^\top z (t - \tau) + g_3(t - \tau) + R_f(t - \tau) + g(t - \tau) \right] \right\} \\ &\quad \times \exp \left\{ \left( \sigma u - \mu(\tau) \right) \left( 1 - \frac{1}{2} (t - \tau)^\top (t - \tau) + C_4(t - \tau) + R_C(t - \tau) \right) \right\} \\ &\quad \times \exp \left\{ \mu(\tau) + \partial \mu(\tau)^\top (t - \tau) + \frac{1}{2} (t - \tau)^\top \Delta \mu(\tau) (t - \tau) + R_\mu(t - \tau) \right\} dt. \end{aligned} \quad (40)$$

where the first row corresponds to the expansion of  $w_t = f_*(t)$ , the second and the third correspond to those of  $C(t)$  and  $\mu(t)$  respectively. We write the exponent inside the integral in a quadratic form of  $(t - \tau)$  and obtain that

$$\begin{aligned} \mathcal{I}(T) &= \exp \left\{ \sigma u + \sigma w + \frac{\sigma}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} \right\} \\ &\quad \times \int_{t \in T} \exp \left\{ -\frac{\sigma}{2} (t - \tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y})^\top (uI - \tilde{\mathbf{z}}) (t - \tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y}) \right\} \\ &\quad \times \exp \{ \sigma g_3(t - \tau) + \sigma u_\tau C_4(t - \tau) + \sigma R(t - \tau) \} \times \exp \{ \sigma g(t - \tau) \} dt, \end{aligned} \quad (41)$$

where  $\tilde{y}$  and  $\tilde{\mathbf{z}}$  are defined as in (24). Let  $a(s)$  and  $b(s)$  be two generic positive functions. Then, we have the representation of the following integral

$$\int_T a(s) b(s) ds = E[b(S)] \int_T a(s) ds$$

where  $S$  is a random variable taking values in  $T$  with density  $a(s)/\int_T a(t) dt$ . Using this representation and the change of variable that  $s = (uI - \tilde{\mathbf{z}})^{1/2} (t - \tau)$ , we write the big integral (41) as a product of expectations and a normalizing constant, and obtain that

$$\begin{aligned} \mathcal{I}(T) &= \det(uI - \tilde{\mathbf{z}})^{-1/2} \exp \{ \sigma u + \sigma w + \frac{\sigma}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} \} \\ &\quad \times \int_{(uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} s + \tau \in T} \exp \left\{ -\frac{\sigma}{2} (s - (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y})^\top (s - (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y}) \right\} ds \\ &\quad \times E \left[ e^{\sigma g_3((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} S) + \sigma u_\tau C_4((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} S) + \sigma R((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} S)} \right] \times E \left[ e^{\sigma g((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} \tilde{S})} \right]. \end{aligned}$$

The two expectations in the above display are taken with respect to  $S$  and  $\tilde{S}$  given the process  $g(t)$ .  $S$  is a random variable taking values in the set  $\{s : (uI - \tilde{\mathbf{z}})^{-1/2} s + \tau \in T\}$  with density proportional to

$$e^{-\frac{\sigma}{2} (s - (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y})^\top (s - (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y})} \quad (42)$$

and  $\tilde{S}$  is a random variable taking values in the set  $\{s : (uI - \tilde{\mathbf{z}})^{-1/2} s + \tau \in T\}$  with density proportional to

$$e^{-\frac{\sigma}{2} (s - (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y})^\top (s - (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y}) + \sigma g_3((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} s) + \sigma u_\tau C_4((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} s) + \sigma R((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} s)}.$$

Together with the definition of  $u$  that  $\left(\frac{2\pi}{\sigma}\right)^{d/2} u^{-d/2} e^{\sigma u} = b$ , we obtain that  $\mathcal{I}(T) > b$  if and only if

$$\begin{aligned} \mathcal{I}(T) &= \det(uI - \tilde{\mathbf{z}})^{-1/2} e^{\sigma u + \sigma w + \frac{\sigma}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y}} \\ &\quad \times \int_{(uI - \mathbf{z})^{-\frac{1}{2}} s + \tau \in T} e^{-\frac{\sigma}{2} (s - (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y})^\top (s - (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y})} ds \\ &\quad \times E \left[ e^{\sigma g_3((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} S) + \sigma u_\tau C_4((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} S) + \sigma R((uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} S)} \right] \cdot e^{-u^{-1} \xi_u} \\ &> \left( \frac{2\pi}{\sigma} \right)^{d/2} u^{-d/2} e^{\sigma u}, \end{aligned} \quad (43)$$

where

$$\xi_u = -u \log \left\{ E \exp \left[ \sigma g \left( (uI - \tilde{\mathbf{z}})^{-\frac{1}{2}} \tilde{S} \right) \right] \right\}. \quad (44)$$

We take log on both sides and plug in the result of Lemma 19 that handles the big expectation term in (43). Then, the inequality (43) is equivalent to

$$\begin{aligned} &w + \frac{1}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} - \frac{1}{2\sigma} \log \det(I - u^{-1} \tilde{\mathbf{z}}) - \frac{1}{8u} (u^{-1} \tilde{Y} + \mathbf{1}/\sigma)^\top \mu_{22} (u^{-1} \tilde{Y} + \mathbf{1}/\sigma) \\ &+ \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2 u} + \frac{1}{8\sigma^2 u} \sum_i \partial_{iii}^4 C(0) + o(u^{-1}) > u^{-1} \sigma^{-1} \xi_u, \end{aligned} \quad (45)$$

where  $\tilde{Y} = \{\tilde{y}_i^2, i = 1, \dots, d; 2\tilde{y}_i \tilde{y}_j, 1 \leq i < j \leq d\}$ , and  $\mathbf{1} = (\underbrace{1, \dots, 1}_d, \underbrace{0, \dots, 0}_{d(d-1)/2})^\top$ . On the set  $\mathcal{L}$ , we further simplify (45) using the following facts (see Lemma 20)

$$\begin{aligned} \partial \mu_\sigma(\tau) &= O(u^{-1/2+\epsilon}), \\ \log \det(I - u^{-1} \tilde{\mathbf{z}}) &= -u^{-1} Tr(\tilde{\mathbf{z}}) + o(u^{-1}) = -\frac{\mathbf{1}^\top (z + \partial^2 \mu_\sigma(\tau)) + d \cdot \mu_\sigma(\tau)}{u} + o(u^{-1}), \end{aligned}$$

where  $Tr$  is the trace of a matrix. Therefore, on the set  $\mathcal{L}$ , (45) is equivalent to

$$w + \frac{y^\top y}{2u} + \frac{\mathbf{1}^\top (z + \partial^2 \mu_\sigma(\tau)) + d \cdot \mu_\sigma(\tau)}{2\sigma u} + \frac{\sum_i \partial_{iii}^4 C(0)}{8\sigma^2 u} + o(u^{-1}) > u^{-1} \sigma^{-1} \xi_u, \quad (46)$$

and further equivalently (by replacing  $u$  with  $u_\tau$ )

$$w + \frac{y^\top y}{2u_\tau} + \frac{\mathbf{1}^\top (z + \partial^2 \mu_\sigma(\tau)) + d \cdot \mu_\sigma(\tau)}{2\sigma u_\tau} + \frac{\sum_i \partial_{iii}^4 C(0)}{8\sigma^2 u_\tau} + o(u^{-1}) > u^{-1} \sigma^{-1} \xi_u.$$

Using the notations defined as in (10) and (32),  $\mathcal{I}(T) > b$  is equivalent to

$$\mathcal{A}_\tau + o(u^{-1}) > u^{-1} \sigma^{-1} \xi_u,$$

where  $\mathcal{A}_\tau$  is defined as in (32).

### 3.3 Part 2

In Part 2, we first consider  $(1 - \rho_1 - \rho_2)K$  in the first expectation of (31) (which is on the set  $\{\mathcal{A}_\tau \geq 0\}$ ) and then  $(1 - \rho_1 - \rho_2)K + \rho_1 K_1$  in the second expectation of (31).

**Part 2.1: the analysis of  $K$  when  $\mathcal{A}_\tau \geq 0$**

Similar to Part 1, all the derivations are conditional on  $(i, \tau, w, y, z)$ . We now proceed to the second part of the proof. More precisely, we simplify the term  $K$  defined as in (29) and write it as a deterministic function of  $(w, y, z)$  with a small correction term. Recall that

$$\begin{aligned} K &= \int_{A^*} l(t) u_t H_\lambda \exp \left\{ -\lambda u_t \left( w_t + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top (\bar{z}_t + \mu_2(t - \tau) u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} - u_t \right) \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20}\mu_{22}^{-1}(\bar{z}_t + \mu_2(t - \tau) u_\tau)|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left| \mu_{22}^{-1/2}(\bar{z}_t + \mu_2(t - \tau) u_\tau) - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right] \right\} \\ &\quad \times h_{x,z}^{-1}(w_t + u_\tau C(t - \tau), z_t + u_\tau \partial^2 C(t - \tau)) dt. \end{aligned}$$

We plug in the form of  $h_{x,z}$  and  $l(t)$  that are defined in (27) and (14) and obtain that

$$\begin{aligned} K &= (2\pi)^{\frac{(d+1)(d+2)}{4} - \frac{d}{2}} \det(\Gamma)^{\frac{1}{2}} \cdot \det(\Delta\mu_\sigma(t_*))^{1/2} u_{t_*}^{d/2} H_\lambda \\ &\quad \times \int_{A^*} \exp \left\{ -\frac{u_{t_*} \cdot (t - t_*)^\top \Delta\mu_\sigma(t_*)(t - t_*)}{2} \right\} \\ &\quad \times u_t \times \exp \left\{ -\lambda u_t \left( w_t + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top (\bar{z}_t + \mu_2(t - \tau) u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} - u_t \right) \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20}\mu_{22}^{-1}(\bar{z}_t + \mu_2(t - \tau) u_\tau)|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left| \mu_{22}^{-1/2}(\bar{z}_t + \mu_2(t - \tau) u_\tau) - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right] \right\} \\ &\quad \times \exp \left\{ \frac{1}{2} \left[ \frac{(w_t + u_\tau C(t - \tau) - \mu_{20}\mu_{22}^{-1}(z_t + \mu_2(t - \tau) u_\tau))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \right. \right. \\ &\quad \left. \left. + (z_t + \mu_2(t - \tau) u_\tau)^\top \mu_{22}^{-1}(z_t + \mu_2(t - \tau) u_\tau) \right] \right\} dt. \end{aligned}$$

For some  $\delta' > \epsilon$ , where  $\epsilon$  is the parameter we used to define  $\mathcal{L}$ , we further restrict the integration region by defining

$$\begin{aligned} \mathcal{I}_2 &= \int_{A^*, |t - \tau| < u^{-1+\delta'}} \exp \left\{ -\frac{u_{t_*} (t - t_*)^\top \Delta\mu_\sigma(t_*)(t - t_*)}{2} \right\} \\ &\quad \times u_t \times \exp \left\{ -\lambda u_t \left( w_t + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top (\bar{z}_t + \mu_2(t - \tau) u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} - u_t \right) \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20}\mu_{22}^{-1}(\bar{z}_t + \mu_2(t - \tau) u_\tau)|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left| \mu_{22}^{-1/2}(\bar{z}_t + \mu_2(t - \tau) u_\tau) - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right] \right\} \\ &\quad \times \exp \left\{ \frac{1}{2} \left[ \frac{(w_t + u_\tau C(t - \tau) - \mu_{20}\mu_{22}^{-1}(z_t + \mu_2(t - \tau) u_\tau))^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} \right. \right. \\ &\quad \left. \left. + (z_t + \mu_2(t - \tau) u_\tau)^\top \mu_{22}^{-1}(z_t + \mu_2(t - \tau) u_\tau) \right] \right\} dt. \end{aligned} \tag{47}$$

Thus,

$$K \geq (2\pi)^{\frac{(d+1)(d+2)}{4} - \frac{d}{2}} \det(\Gamma)^{\frac{1}{2}} \cdot \det(\Delta\mu_\sigma(t_*))^{1/2} u_{t_*}^{d/2} H_\lambda \cdot \mathcal{I}_2.$$



For the rest of Part 2.1, we focus on  $\mathcal{I}_2$ . With some tedious algebra, Lemma 21 writes  $\mathcal{I}_2$  in a more manageable form, that is,  $\mathcal{I}_2$  equals

$$\begin{aligned} & \int_{A^*, |t-\tau| < u^{-1+\delta'}} \exp \left\{ -\frac{u_{t_*}(t-t_*)^\top \Delta \mu_\sigma(t_*)(t-t_*)}{2} + \frac{u_t^2}{2} \right\} \times u_t \\ & \times \exp \left\{ (1-\lambda)u_t [w_t + u_\tau C(t-\tau) - u_t] + (1-\lambda) \frac{\mathbf{1}^\top (z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau)}{2\sigma} - \lambda B_t - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ & \times \exp \left\{ \frac{(w_t + u_\tau C(t-\tau) - u_t)^2 - 2(w_t + u_\tau C(t-\tau) - u_t) \mu_{20} \mu_{22}^{-1} (z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau)}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} \right\} dt. \end{aligned} \quad (48)$$

Lemma 22 implies that  $\{|t-\tau| < u^{-1+\delta'}\} \subset A^*$ . Thus, on the set  $\{\mathcal{A}_\tau > 0\}$ , we have  $A^* \cap \{|t-\tau| < u^{-1+\delta'}\} = \{|t-\tau| < u^{-1+\delta'}\}$  and we can remove  $A^*$  from the integration region of  $\mathcal{I}_2$ . In addition, on the set  $\mathcal{L}$  and  $|t-\tau| < u^{-1+\delta'}$ , we have that

$$\begin{aligned} u_\tau - u_t C(t-\tau) &= O(u^{-1+2\delta'}), \quad \mu_2(t-\tau) = \mu_{20} + O(|t-\tau|^2), \\ |u_\tau \mu_2(t-\tau) - u_t \mu_{20}| &= O(u^{-1+2\delta'}), \quad (u_\tau - u_t C(t-\tau))|z_t| = o(1). \end{aligned}$$

We insert the above estimates to (48). Together with the fact that

$$\exp \left\{ -\frac{u_{t_*}(t-t_*)^\top \Delta \mu_\sigma(t_*)(t-t_*)}{2} + \frac{u_t^2}{2} \right\} = (1 + o(1)) \exp \left\{ \frac{1}{2} u_{t_*}^2 \right\},$$

we have that

$$\begin{aligned} \mathcal{I}_2 &\sim u \times \exp \left\{ \frac{1}{2} u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ &\times \int_{|t-\tau| < u^{-1+\delta'}} \exp \left\{ (1-\lambda)u_t [w_t + u_\tau (C(t-\tau) - 1) + (\mu_\sigma(t) - \mu_\sigma(\tau))] + (1-\lambda) \frac{\mathbf{1}^\top z}{2\sigma} \right\} \\ &\times \exp \left\{ \frac{w_t^2 - 2w_t \mu_{20} \mu_{22}^{-1} z_t + o(1)w_t}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} \right\} dt. \end{aligned}$$

Further, we have that

$$w_t^2 - 2w_t \mu_{20} \mu_{22}^{-1} z_t + o(1)w_t = o(1) + u \cdot w \cdot O(u^{-1/2+\delta'}).$$

Let  $\zeta_u = O(u^{-1/2+\delta'})$  and we simplify  $\mathcal{I}_2$  to

$$\begin{aligned} \mathcal{I}_2 &\sim u \times \exp \left\{ \frac{1}{2} u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ &\times \int_{|t-\tau| < u^{-1+\delta'}} \exp \left\{ (1-\lambda)(u_\tau + \zeta_u) \left[ \zeta_u w + w_t + u_\tau (C(t-\tau) - 1) + (\mu_\sigma(t) - \mu_\sigma(\tau)) \right] \right. \\ &\quad \left. + (1-\lambda) \frac{\mathbf{1}^\top z}{2\sigma} \right\} dt \end{aligned}$$

In what follows, we insert the expansions in (34), (35) and (36) into the expression of  $\mathcal{I}_2$  and write the exponent as a quadratic function of  $t - \tau$ , and obtain that on the set  $\mathcal{L}$

$$\begin{aligned} \mathcal{I}_2 \sim & u \times \exp \left\{ \frac{1}{2} u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ & \times \exp \left\{ (1 - \lambda)(u_\tau + \zeta_u) \left( (1 + \zeta_u)w + \frac{1}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} \right) \right\} \\ & \times \int_{|t-\tau| < u^{-1+\delta'}} \exp \left\{ -\frac{(1 - \lambda)(u_\tau + \zeta_u)}{2} (t - \tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y})^\top (uI - \tilde{\mathbf{z}}) (t - \tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y}) \right\} \\ & \times \exp \{ (1 - \lambda)(u_\tau + \zeta_u) [g_3(t - \tau) + u_\tau C_4(t - \tau) + R(t - \tau) + g(t - \tau)] \} dt, \end{aligned} \quad (49)$$

where we recall that  $\tilde{y} = y + \partial \mu_\sigma(\tau)$ , and  $\tilde{\mathbf{z}} = \mathbf{z} + u_\sigma(\tau)I + \Delta \mu_\sigma(\tau)$ . This derivation is very similar to that from (40) to (41). In the last row of the above display, on the set  $\mathcal{L}$  and  $|t - \tau| < u^{-1+\delta'}$ ,

$$u g_3(t - \tau) + u^2 C_4(t - \tau) + u R(t - \tau) = o(1).$$

Therefore, they can be ignored. We consider the change of variable that

$$s = (1 - \lambda)^{1/2} (u_\tau + \zeta_u)^{1/2} (uI - \tilde{\mathbf{z}})^{1/2} (t - \tau)$$

and obtain that  $\mathcal{I}_2$  equals (with the terms  $g_3$  and  $C_4$  removed)

$$\begin{aligned} \mathcal{I}_2 \sim & (1 - \lambda)^{-d/2} u^{-d+1} \exp \left\{ \frac{1}{2} u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ & \times \exp \left\{ (1 - \lambda)(u_\tau + \zeta_u) \left( (1 + \zeta_u)w + \frac{1}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} + \frac{\mathbf{1}^\top z}{2\sigma u} \right) \right\} \\ & \times \int_{s \in \mathcal{S}_u} e^{-\frac{1}{2} |s - (1 - \lambda)^{1/2} (u_\tau + \zeta_u)^{1/2} (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y}|^2} ds \\ & \times E \left[ e^{(1 - \lambda)(u_\tau + \zeta_u)g((1 - \lambda)^{-1/2} (u_\tau + \zeta_u)^{-1/2} (uI - \tilde{\mathbf{z}})^{-1/2} S')} \right], \end{aligned} \quad (50)$$

where  $\mathcal{S}_u = \{s : |(1 - \lambda)^{-1/2} (u_\tau + \zeta_u)^{-1/2} (uI - \tilde{\mathbf{z}})^{-1/2} s| < u^{-1+\delta'}\}$  and  $S'$  is a random variable taking values on the set  $\mathcal{S}_u$  with density proportional to

$$e^{-\frac{1}{2} |s - (1 - \lambda)^{1/2} (u_\tau + \zeta_u)^{1/2} (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y}|^2}.$$

We use  $\kappa$  to denote the last two terms of (50)

$$\begin{aligned} \kappa = & \int_{\mathcal{S}_u} e^{-\frac{1}{2} |s - (1 - \lambda)^{1/2} (u_\tau + \zeta_u)^{1/2} (uI - \tilde{\mathbf{z}})^{-1/2} \tilde{y}|^2} ds \\ & \times E \left[ e^{(1 - \lambda)(u_\tau + \zeta_u)g((1 - \lambda)^{-1/2} (u_\tau + \zeta_u)^{-1/2} (uI - \tilde{\mathbf{z}})^{-1/2} S')} \right]. \end{aligned} \quad (51)$$

It is helpful to keep in mind that  $\kappa$  is approximately  $(2\pi)^{d/2}$ . We insert  $\kappa$  back to the expression of  $\mathcal{I}_2$ . Together with the fact that  $\tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} = |\tilde{y}|^2/u + o(u^{-1})$ , we have

$$\begin{aligned} \mathcal{I}_2 \sim & \kappa (1 - \lambda)^{-d/2} u^{-d+1} \exp \left\{ \frac{1}{2} u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ & \times \exp \left\{ (1 - \lambda)(u_\tau + \zeta_u) \left( (1 + \zeta_u)w + \frac{|\tilde{y}|^2}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} \right) \right\}. \end{aligned} \quad (52)$$

Thus, we have that on the set  $\{\mathcal{A}_\tau > 0\}$ ,

$$\begin{aligned} K &\geq (2\pi)^{\frac{(d+1)(d+2)}{4}-\frac{d}{2}} \det(\Gamma)^{\frac{1}{2}} \cdot \det(\Delta\mu_\sigma(t_*))^{1/2} u_{t_*}^{d/2} H_\lambda \cdot \mathcal{I}_2 \\ &= (\kappa + o(1))(2\pi)^{\frac{(d+1)(d+2)}{4}-\frac{d}{2}} \det(\Gamma)^{\frac{1}{2}} \cdot \det(\Delta\mu_\sigma(t_*))^{1/2} H_\lambda \cdot (1-\lambda)^{-d/2} u^{-d/2+1} \\ &\quad \times \exp \left\{ \frac{1}{2} u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} + (1-\lambda)(u_\tau + \zeta_u) \left( (1+\zeta_u)w + \frac{|\tilde{y}|^2}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} \right) \right\}. \end{aligned} \quad (53)$$

We further insert the  $\mathcal{A}_\tau$  defined in (32) into (53) and obtain that

$$\begin{aligned} K &\geq (\kappa + o(1))(2\pi)^{\frac{(d+1)(d+2)}{4}-\frac{d}{2}} \det(\Gamma)^{\frac{1}{2}} \cdot \det(\Delta\mu_\sigma(t_*))^{1/2} H_\lambda \cdot (1-\lambda)^{-d/2} u^{-d/2+1} \\ &\quad \times \exp \left\{ \frac{1}{2} u_{t_*}^2 - B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} + (1-\lambda)u_\tau(1+o(1))\mathcal{A}_\tau + (1-\lambda)\zeta_u \cdot (|\tilde{y}|^2 + |z|) \right\}. \end{aligned} \quad (54)$$

## Part 2.2: the analysis of $dP/dQ$ when $\mathcal{A}_\tau < 0$

In this part, we focus mostly on  $K_1$  term, whose handling is very similar to that of  $K$ . Therefore, we only list out the key steps. For some large constant  $M$ , let

$$D = \{|t - \tau - (uI - \tilde{\mathbf{z}})^{-1}\tilde{y}| < Mu^{-1}\}$$

that is the dominating region of the integral. We split the set  $D = (A^* \cap D) \cup ((A^*)^c \cap D)$ . There are two situations:  $\text{mes}((A^*)^c \cap D) > \text{mes}(A^* \cap D)$  and  $\text{mes}((A^*)^c \cap D) \leq \text{mes}(A^* \cap D)$ . For the first situation, the term  $K_1$  is dominating; for the second situation, the term  $K$  (more precisely  $\mathcal{I}_2$ ) is dominating.

To simplify  $K_1$ , we write it as

$$\begin{aligned} K_1 &= (2\pi)^{\frac{(d+1)(d+2)}{4}-\frac{d}{2}} \det(\Gamma)^{\frac{1}{2}} \cdot \det(\Delta\mu_\sigma(t_*))^{1/2} u_{t_*}^{d/2} H_{\lambda_1} \times \left[ \int_{(A^*)^c \cap D} \dots + \int_{(A^*)^c \cap D^c} \dots \right] \\ &\triangleq (2\pi)^{\frac{(d+1)(d+2)}{4}-\frac{d}{2}} \det(\Gamma)^{\frac{1}{2}} \cdot \det(\Delta\mu_\sigma(t_*))^{1/2} u_{t_*}^{d/2} H_{\lambda_1} \times [\mathcal{I}_{1,2} + \mathcal{I}_{1,3}]. \end{aligned}$$

Note that the difference between  $K_1$  and  $K$  is that the term “ $-\lambda$ ” has been replace by “ $\lambda_1$ ”. With exactly the same derivation for (48), we obtain that  $\mathcal{I}_{1,2}$  equals (by replacing “ $-\lambda$ ” in (48) by “ $\lambda_1$ ”)

$$\begin{aligned} &\int_{(A^*)^c \cap D} \exp \left\{ -\frac{u_{t_*}(t-t_*)^\top \Delta\mu_\sigma(t_*)(t-t_*)}{2} + \frac{1}{2} u_t^2 \right\} \times u_t \\ &\times \exp \left\{ (1+\lambda_1)u_t [w_t + u_\tau C(t-\tau) - u_t] + (1+\lambda_1) \frac{\mathbf{1}^\top (z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau)}{2\sigma} + \lambda_1 B_t - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ &\times \exp \left\{ \frac{(w_t + u_\tau C(t-\tau) - u_t)^2 - 2(w_t + (u_\tau C(t-\tau) - u_t)) \mu_{20} \mu_{22}^{-1} (z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau)}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} \right\} dt. \end{aligned} \quad (55)$$

With a very similar derivation as in Part 2.1, in particular, the result in (49), we have that

$$\begin{aligned} \mathcal{I}_{1,2} &\sim u \exp \left\{ \frac{1}{2} u_{t_*}^2 + \lambda_1 B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ &\times \exp \left\{ (1+\lambda_1)(u_\tau + \zeta_u) \left( (1+\zeta_u)w + \frac{1}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} + \frac{\mathbf{1}^\top z}{2\sigma u} \right) \right\} \\ &\times \int_{(A^*)^c \cap D} \exp \left\{ (1+\lambda_1)(u_\tau + \zeta_u) \left[ -\frac{1}{2} (t-\tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y})^\top (uI - \tilde{\mathbf{z}}) (t-\tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y}) \right] \right\} \\ &\times \exp \{ (1+\lambda_1)(u_\tau + \zeta_u) [g_3(t-\tau) + u_\tau C_4(t-\tau) + R(t-\tau) + g(t-\tau)] \} dt. \end{aligned} \quad (56)$$

Furthermore, similar to the results in (52), we have that

$$\begin{aligned}\mathcal{I}_{1,2} &= (1 + o(1))\kappa_{1,2}(1 + \lambda_1)^{-d/2}u^{-d+1}e^{\frac{1}{2}u_{t_*}^2 + \lambda_1 B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2}} \\ &\quad \times e^{(1+\lambda_1)(u_\tau + \zeta_u) \left( (1+\zeta_u)w + \frac{1}{2}\tilde{y}^\top (uI - \tilde{\mathbf{Z}})^{-1}\tilde{y} + \frac{\mathbf{1}^\top \mathbf{z}}{2\sigma u_\tau} \right)},\end{aligned}\quad (57)$$

where

$$\begin{aligned}\kappa_{1,2} &= \int_{t(s) \in (A^*)^c \cap D} e^{-\frac{1}{2}|s - (1+\lambda_1)^{1/2}(u_\tau + \zeta_u)^{1/2}(uI - \tilde{\mathbf{Z}})^{-1/2}\tilde{y}|^2} ds \\ &\quad \times E \left[ e^{(1+\lambda_1)(u_\tau + \zeta_u)g((1+\lambda_1)^{-1/2}(u_\tau + \zeta_u)^{-1/2}(uI - \tilde{\mathbf{Z}})^{-1/2}S_{1,2})} \right],\end{aligned}$$

the change of variable  $t(s) = (1 - \lambda)^{-1/2}(u_\tau + \zeta_u)^{-1/2}(uI - \tilde{\mathbf{Z}})^{-1/2}s$ ,  $S_{1,2}$  is a random variable taking values in the set  $\{s : t(s) \in (A^*)^c \cap D\}$  with an appropriately chosen density function similarly as in (50). In summary, the only difference between  $\mathcal{I}_{1,2}$  and  $\mathcal{I}_2$  lies in that the multiplier  $-\lambda$  is replaced by  $\lambda_1$ .

We now proceed to providing a lower bound of  $(1 - \rho_1 - \rho_2)K + \rho_1 K_1$ . Note that

$$\max\{\text{mes}((A^*)^c \cap D), \text{mes}(A^* \cap D)\} \geq \frac{1}{2}\text{mes}(D).$$

Therefore at least one of  $(A^*)^c \cap D$  and  $A^* \cap D$  is nonempty. If  $\text{mes}((A^*)^c \cap D) \geq \frac{1}{2}\text{mes}(D)$ , we have the bound

$$(1 - \rho_1 - \rho_2)K + \rho_1 K_1 \geq \rho_1 K_1 \geq \Theta(1)\rho_1 u^{d/2}\mathcal{I}_{1,2}.$$

Similarly, if  $\text{mes}(A^* \cap D) \geq \frac{1}{2}\text{mes}(D)$ , we have that

$$(1 - \rho_1 - \rho_2)K + \rho_1 K_1 \geq \Theta(1)(1 - \rho_1 - \rho_2)u^{d/2}\mathcal{I}_2.$$

We further split  $\mathcal{I}_2$  in Part 2.1 into two parts:

$$\begin{aligned}\mathcal{I}_2 &= \int_{A^* \cap D} \cdots dt + \int_{A^* \cap D^c} \cdots dt \\ &\triangleq \mathcal{I}_{2,1} + \mathcal{I}_{2,2}.\end{aligned}\quad (58)$$

Similar to the derivation of  $\mathcal{I}_{1,2}$ , we have that

$$\begin{aligned}\mathcal{I}_{2,1} &\sim \kappa_{2,1}(1 - \lambda)^{-d/2}u^{-d+1}e^{\frac{1}{2}u_{t_*}^2 - \lambda B_{t_*} - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2}} \\ &\quad \times e^{(1-\lambda)(u_\tau + \zeta_u) \left( (1+\zeta_u)w + \frac{|\tilde{y}|^2}{2u_\tau} + \frac{\mathbf{1}^\top \mathbf{z}}{2\sigma u_\tau} \right)},\end{aligned}\quad (59)$$

where

$$\begin{aligned}\kappa_{2,1} &= \int_{t(s) \in A^* \cap D} e^{-\frac{1}{2}|s - (1-\lambda)^{1/2}(u_\tau + \zeta_u)^{1/2}(uI - \tilde{\mathbf{Z}})^{-1/2}\tilde{y}|^2} ds \\ &\quad \times E \left[ e^{(1-\lambda)(u_\tau + \zeta_u)g((1-\lambda)^{-1/2}(u_\tau + \zeta_u)^{-1/2}(uI - \tilde{\mathbf{Z}})^{-1/2}S_{2,1})} \right].\end{aligned}\quad (60)$$

$S_{2,1}$  is a random variable taking values on the set  $\{s : t(s) \in A^* \cap D\}$  with an appropriate density function similarly as in (50).

Then combining the above results of  $\mathcal{I}_{1,2}$  and  $\mathcal{I}_{2,1}$ , we have that for the case in which  $\mathcal{A}_\tau < 0$

$$\begin{aligned} \rho_1 K_1 + (1 - \rho_1 - \rho_2)K &\geq \Theta(1)u^{d/2} [\mathbb{I}_{C_1} \rho_1 \mathcal{I}_{1,2} + \mathbb{I}_{C_2} (1 - \rho_1 - \rho_2) \mathcal{I}_{2,1}] \\ &\geq \Theta(1)u^{-d/2+1} e^{\frac{1}{2}u_{t_*}^2} \\ &\quad \times \left[ \mathbb{I}_{C_1} \cdot \rho_1 \kappa_{1,2} e^{(1+\lambda_1)(u_\tau + \zeta_u) \left( (1+\zeta_u)w + \frac{|\tilde{y}|^2}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} \right)} \right. \\ &\quad \left. + \mathbb{I}_{C_2} \cdot (1 - \rho_1 - \rho_2) \kappa_{2,1} e^{(1-\lambda)(u_\tau + \zeta_u) \left( (1+\zeta_u)w + \frac{|\tilde{y}|^2}{2u_\tau} + \frac{\mathbf{1}^\top z}{2\sigma u_\tau} \right)} \right] \end{aligned}$$

where  $C_1 = \{f(\cdot) : \text{mes}((A^*)^c \cap D) \geq \text{mes}(A^* \cap D)\}$  and  $C_2 = C_1^c$ . We further insert  $\mathcal{A}_\tau$  defined in (32). Note that on the set  $\{\mathcal{A}_\tau < 0\}$ ,  $(1 + \lambda_1)\mathcal{A}_\tau < (1 - \lambda)\mathcal{A}_\tau$  and  $B_t$  is bounded away from zero and infinity. Then,

$$\begin{aligned} (1 - \rho_1 - \rho_2)K + \rho_1 K_1 &\geq \Theta(1)u^{-d/2+1} e^{\frac{1}{2}u_{t_*}^2} \cdot e^{(1+\lambda_1)(1+\zeta_u)u_\tau \mathcal{A}_\tau + \zeta_u \cdot (|\tilde{y}|^2 + |z|)} \\ &\quad \times \left[ \mathbb{I}_{C_1} \cdot \rho_1 \kappa_{1,2} + \mathbb{I}_{C_2} \cdot (1 - \rho_1 - \rho_2) \kappa_{2,1} \right]. \end{aligned} \quad (61)$$

### Part 3

We now put together the results in Part 1 and Part 2 and obtain an approximation for (31). Recall that

$$\begin{aligned} E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{E}_b, \mathcal{L} \right] &\leq E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \right] \\ &\quad + E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0 \right]. \end{aligned} \quad (62)$$

We consider the two terms on the right-hand-side of the above display one by one. We start with the first term

$$\begin{aligned} E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \right] &= E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0, \iota = 0 \right] \\ &\quad + E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K]^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0, \iota = 1 \right]. \end{aligned} \quad (63)$$

The index  $\tau$  admits density  $l(t)$  when  $\iota = 0$  and  $\tau$  is uniformly distributed over  $T$  if  $\iota = 1$ .

Consider the first expectation in (63). Note that conditional on  $\tau$  and  $\iota = 0$ , on the set  $\mathcal{L} \cap \{\mathcal{A}_\tau \geq 0\}$ ,  $(w, y, z)$  follows density  $(1 - \rho_1 - \rho_2)h_{0,\tau}^*(w, y, z)/(1 - \rho_2)$  defined as in (25). Thus, according to (54), we have that the conditional expectation

$$\begin{aligned} &E^Q \left[ \frac{1}{(1 - \rho_1 - \rho_2)^2 K^2} ; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \middle| \iota = 0, \tau \right] \\ &\leq (1 + o(1)) \left[ \frac{H_\lambda^{-1} \det(\Gamma)^{-\frac{1}{2}} \det(\Delta \mu_\sigma(t_*))^{-1/2}}{(2\pi)^{\frac{(d+1)(d+2)}{4} - \frac{d}{2}}} (1 - \lambda)^{d/2} u^{d/2-1} e^{-\frac{1}{2}u_{t_*}^2 + B_{t_*} + \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2}} \right]^2 \\ &\quad \times \int_{A_\tau > 0, \mathcal{L}} e^{-2(1-\lambda)u \left( (1+o(1))\mathcal{A}_\tau + o(1)\frac{|\tilde{y}|^2}{2u} + o(1)\frac{\mathbf{1}^\top z}{2\sigma u} \right)} \times \gamma_u(u\sigma \mathcal{A}_\tau) \times \frac{1 - \rho_1 - \rho_2}{1 - \rho_2} h_{0,\tau}^*(w, y, z) dw dy dz \end{aligned} \quad (64)$$

where

$$\gamma_u(x) = E \left[ \frac{1}{(1 - \rho_1 - \rho_2)^2 \kappa^2}; x > \xi_u \middle| \imath, \tau, w, y, z \right],$$

with the expectation taken with respect to the process  $g(t)$ . We insert the analytic form of  $h_{0,\tau}^*(w, y, z)$  in (25) and obtain that

$$\begin{aligned} & \int_{\mathcal{A}_\tau > 0, \mathcal{L}} e^{-2(1-\lambda)u \left( (1+o(1))\mathcal{A}_\tau + o(1)\frac{|y|^2}{2u} + o(1)\frac{\mathbf{1}^\top z}{2\sigma u} \right)} \times \gamma_u(u\sigma\mathcal{A}_\tau) \times \frac{1 - \rho_1 - \rho_2}{1 - \rho_2} h_{0,\tau}^*(w, y, z) dw dy dz \\ &= \frac{(1 - \rho_1 - \rho_2)H_\lambda \cdot u_\tau}{1 - \rho_2} \int_{\mathcal{A}_\tau > 0} \gamma_u(u\sigma\mathcal{A}_\tau) \exp \left\{ -2(1 - \lambda + o(1))u\mathcal{A}_\tau + o(|z| + |y|^2) \right\} \\ & \quad \times \exp \left\{ -\lambda u_\tau \mathcal{A}_\tau - \frac{1}{2} \left[ \frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left| \mu_{22}^{-1/2}z - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right] - \frac{1 - \lambda}{2} y^\top y \right\} d\mathcal{A}_\tau dy dz. \end{aligned} \quad (65)$$

Thanks to the Borel-TIS inequality (Lemma 15), Lemma 18 and the definition of  $\kappa$  in (51), for  $x > 0$ ,  $\gamma_u(x)$  is bounded and as  $b \rightarrow \infty$ ,

$$E \left[ \frac{1}{\kappa^2}; x > \xi_u \right] \rightarrow (2\pi)^{-d}.$$

Thus, by the dominated convergence theorem and with  $H_\lambda$  defined as in (15), as  $u \rightarrow \infty$ , we have that

$$(65) \sim \frac{(2\pi)^{-d}}{(1 - \rho_1 - \rho_2)(1 - \rho_2)} \frac{e^{-\lambda\eta\lambda}}{2 - \lambda}.$$

We insert it back to (64) and obtain that

$$\begin{aligned} & E^Q \left[ \frac{1}{(1 - \rho_1 - \rho_2)^2 K^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \middle| \imath = 0, \tau \right] \\ & \leq (1 + o(1)) \frac{(2\pi)^{-d}}{(1 - \rho_1 - \rho_2)(1 - \rho_2)} \frac{e^{-\lambda\eta\lambda}}{2 - \lambda} \\ & \quad \times \left[ \frac{H_\lambda^{-1} \det(\Gamma)^{-\frac{1}{2}} \det(\Delta\mu_\sigma(t_*))^{-1/2}}{(2\pi)^{\frac{(d+1)(d+2)}{4} - \frac{d}{2}}} (1 - \lambda)^{d/2} u^{d/2-1} e^{-\frac{1}{2}u_{t_*}^2 + B_{t_*} + \frac{\mathbf{1}^\top \mu_{22}\mathbf{1}}{8\sigma^2}} \right]^2. \end{aligned} \quad (66)$$

Using the asymptotic approximation of  $v(b)$  given by Proposition 13, we obtain that

$$E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0, \imath = 0 \right] \leq \frac{1 + o(1)}{1 - \rho_1 - \rho_2} \frac{e^{\lambda\eta}}{\lambda(2 - \lambda)} v^2(b). \quad (67)$$

Note that as  $b$  goes to infinity,  $\rho_1, \rho_2$ , and  $\eta$  are positive converging to 0 and  $\lambda$  is smaller than and converging to 1. Then, the right-hand-side of the above inequality is bounded by  $(1 + \varepsilon)v^2(b)$ .

The handling of the second term of (63) is similar except that  $(w, y, z)$  follows density  $h_\tau^*(w, y, z)$ .



Thus, we only mention the key steps. Note that

$$\begin{aligned}
& E^Q \left[ \frac{1}{(1 - \rho_1 - \rho_2)^2 K^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0 \middle| \iota = 1, \tau \right] \\
&= (1 + o(1)) \left[ \frac{H_\lambda^{-1} \det(\Gamma)^{-\frac{1}{2}} \det(\Delta \mu_\sigma(t_*))^{-1/2}}{(2\pi)^{\frac{(d+1)(d+2)}{4} - \frac{d}{2}}} (1 - \lambda)^{d/2} u^{d/2-1} e^{-\frac{1}{2}u_{t_*}^2 + B_{t_*} + \frac{1^\top \mu_{22}^{-1} 1}{8\sigma^2}} \right]^2 \\
&\quad \times \frac{\det(\Gamma)^{-\frac{1}{2}}}{(2\pi)^{\frac{(d+1)(d+2)}{4}}} \int_{\mathcal{A}_\tau \geq 0, \mathcal{L}} \gamma_u(u\sigma \mathcal{A}_\tau) e^{-2(1-\lambda)u\mathcal{A}_\tau} e^{-\frac{1+o(1)}{2} \left[ y^\top y + \frac{|w - \mu_{20} \mu_{22}^{-1} z|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + z^\top \mu_{22}^{-1} z \right]} d\mathcal{A}_\tau dy dz \\
&= O(1)(1 - \lambda)^{-1} u^{-1} \cdot u^{d-2} e^{-u_{t_*}^2}. \tag{68}
\end{aligned}$$

According to the asymptotic form of  $v(b)$ , we have that for properly chosen  $\rho_2$  and  $\lambda$  (e.g.  $\rho_2 = 1 - \lambda = 1/\log u$ )

$$E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \geq 0, \iota = 1 \right] = O(1)\rho_2(1 - \lambda)^{-1} u^{-1} \cdot u^{d-2} e^{-u_{t_*}^2} = o(1)v^2(b). \tag{69}$$

Therefore, combining the results in (67) and (69), we have the first term in (62) is bounded by  $(1 + 2\varepsilon)v^2(b)$ .

The last step is to show that the second term of (62) is of a smaller order of  $v^2(b)$ . First, we split the expectation

$$\begin{aligned}
& E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0 \right] \\
&= E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0, \iota = 1 \right] \\
&\quad + E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, -\eta/u_\tau < \mathcal{A}_\tau < 0, \iota = 0 \right] \\
&\quad + E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \leq -\eta/u_\tau, \iota = 0 \right]. \tag{70}
\end{aligned}$$

We study these three terms one by one. Let

$$\gamma_{1,u}(x) = E \left[ \frac{1}{[\mathbb{I}_{C_1} \cdot \rho_1 \kappa_{1,2} + \mathbb{I}_{C_2} \cdot (1 - \rho_1 - \rho_2) \kappa_{2,1}]^2}; x > \xi_u \middle| \iota, \tau, w, y, z \right] \tag{71}$$

We start with the first expectation in (70). Plugging in the lower bound for  $(1 - \rho_1 - \rho_2)K + \rho_1 K_1$  derived in (61), we have

$$\begin{aligned}
& E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0 \middle| \iota = 1, \tau \right] \\
&= O(1)u^{d-2} e^{-u_{t_*}^2} \times \int_{\mathcal{A}_\tau < 0} \gamma_{1,u}(u\sigma \mathcal{A}_\tau) e^{-2(1+\lambda_1)u\mathcal{A}_\tau} e^{-\frac{1}{2} \left[ y^\top y + \frac{|w - \mu_{20} \mu_{22}^{-1} z|^2}{1 - \mu_{20} \mu_{22}^{-1} \mu_{02}} + z^\top \mu_{22}^{-1} z \right]} d\mathcal{A}_\tau dy dz. \tag{72}
\end{aligned}$$

We deal with the  $\gamma_{1,u}(u\sigma \mathcal{A}_\tau)$  term in the above integration. On the set  $\mathcal{L}$ ,  $u\sigma \mathcal{A}_\tau > -u^{3/2+\epsilon}$ . By Lemma 23, for  $-u^{3/2+\epsilon} < x < 0$ , there exists a constant  $\delta^* > 0$  such that

$$E \left[ \frac{1}{\rho_1^2 \kappa_{1,2}^2}; x > \xi_u \middle| \iota, \tau, w, y, z, C_1 \right] = O(1)\rho_1^{-2} e^{u^{\delta^*} x}.$$

and

$$E \left[ \frac{1}{(1 - \rho_1 - \rho_2)^2 \kappa_{2,1}^2}; x > \xi_u \middle| \iota, \tau, w, y, z, C2 \right] = O(1)(1 - \rho_1 - \rho_2)^{-2}(1 - \lambda)^{-d} e^{u^{\delta^*} x}.$$

Therefore, the above approximations and the dominated convergence theorem imply that conditional on  $\mathcal{L}$ ,

$$\int_{\mathcal{A}_\tau < 0} \gamma_{1,u}(u\sigma\mathcal{A}_\tau) e^{-2(1+\lambda_1)u\mathcal{A}_\tau} d\mathcal{A}_\tau = O(1) \cdot \max\{\rho_1^{-2}, (1 - \lambda)^{-d}\} \cdot u^{-1-\delta^*}.$$

Thus, (72) equals

$$(72) = O(1) \max\{\rho_1^{-2}, (1 - \lambda)^{-d}\} \cdot u^{-1-\delta^*} \cdot u^{d-2} e^{-u_{i_*}^2}.$$

Taking expectation of the above equation with respect to  $\iota$  and  $\tau$  and properly choosing the convergence rate of  $\rho_1, \rho_2$  and  $1 - \lambda$ , we have

$$E \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau < 0, \iota = 1 \right] = o(1) \cdot u^{d-2} e^{-u_{i_*}^2} = o(1)v^2(b). \quad (73)$$

For the second term in (70), with the same bound of  $\gamma_{1,u}$  and carefully chosen parameters, we have

$$\begin{aligned} & E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, -\eta/u_\tau < \mathcal{A}_\tau < 0 \middle| \iota = 0, \tau \right] \\ &= O(1) u^{d-2} e^{-u_{i_*}^2} \times u_\tau \int_{-\frac{\eta}{u_\tau} < \mathcal{A}_\tau < 0} \gamma_{1,u}(u\sigma\mathcal{A}_\tau) e^{-2(1+\lambda_1)u\mathcal{A}_\tau} e^{-\lambda u_\tau \mathcal{A}_\tau} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left[ \frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left| \mu_{22}^{-1/2}z - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right] - \frac{1-\lambda}{2} y^\top y \right\} d\mathcal{A}_\tau dy dz \\ &= O(1) \cdot \max\{\rho_1^{-2}, (1 - \lambda)^{-d}\} \cdot u^{-\delta^*} \cdot u^{d-2} e^{-u_{i_*}^2} \\ &= o(1)v^2(b). \end{aligned}$$

and similarly for the third term in (70),

$$\begin{aligned} & E^Q \left[ \frac{1}{[(1 - \rho_1 - \rho_2)K + \rho_1 K_1]^2}; \mathcal{E}_b, \mathcal{L}, \mathcal{A}_\tau \leq -\eta/u_\tau \middle| \iota = 0, \tau \right] \\ &= O(1) \rho_1 \cdot u^{d-2} e^{-u_{i_*}^2} \\ &\quad \times u_\tau \int_{\mathcal{A}_\tau < -\frac{\eta}{u_\tau}} \gamma_{1,u}(u\sigma\mathcal{A}_\tau) e^{-2(1+\lambda_1)u\mathcal{A}_\tau} e^{\lambda_1 u_\tau \mathcal{A}_\tau - \frac{1}{2} \left[ \frac{|\mu_{20}\mu_{22}^{-1}z|^2}{1 - \mu_{20}\mu_{22}^{-1}\mu_{02}} + \left| \mu_{22}^{-1/2}z - \frac{\mu_{22}^{1/2}\mathbf{1}}{2\sigma} \right|^2 \right] - \frac{1+\lambda_1}{2} y^\top y} d\mathcal{A}_\tau dy dz \\ &= O(1) \rho_1 \cdot \max\{\rho_1^{-2}, (1 - \lambda)^{-d}\} \cdot u^{-\delta^*} \cdot u^{d-2} e^{-u_{i_*}^2} \\ &= o(1)v^2(b). \end{aligned} \quad (74)$$

Thus, we put all the estimates in (67), (69), (73), and (74) back to (62). Then, for any  $\varepsilon > 0$ , we can choose  $\rho_1, \rho_2, \eta$  small and  $\lambda$  close to 1 (e.g.,  $\eta = \rho_1 = \rho_2 = 1 - \lambda = 1/\log u$ ) such that for  $b$  sufficiently large

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2; \mathcal{E}_b, \mathcal{L} \right] \leq (1 + 3\varepsilon)v^2(b).$$

We complete the proof of Theorem 3 for the case that  $\mu(t) \neq 0$ .

### 3.4 Case 2: constant mean function.

The proof when  $\mu(t) \equiv 0$  is very similar, except that we need to consider two situations: first,  $\tau$  is not close to the boundary of  $T$  and otherwise. More precisely, for a given  $\delta' > 0$  small enough, we consider the case when  $\tau \in \{t : |t - \tau| \leq u^{-1/2+\delta'}\} \subset T$  and otherwise.

For the first situation,  $\tau$  is “far away” from the boundary of  $T$ , which is the important case, the derivation is same as that of the case where  $\mu(t)$  is not a constant. For the case in which  $\tau$  is within  $u^{-1/2+\delta'}$  distance from the boundary of  $T$ , the contribution of the boundary case is  $o(v^2(b))$ . An intuitive interpretation is that the important region of the integral  $\mathcal{I}(T)$  might be cut off by the boundary of  $T$ . Therefore, in cases that  $\tau$  is too close to the boundary, the tail  $\mathcal{I}(T)$  is not heavier than that of the interior case. The rigorous analysis is basically repeating the Parts 1, 2, and 3 on a truncated region. Therefore, we omit the details.

## 4 Proof of Theorem 4

The proof of Theorem 4 is analogous to that of Theorem 3. According to Lemma 17, we focus on the set  $\mathcal{L}$ . A similar three-part procedure is applied here.

In Part 1, using the transformation from  $f$  to the process  $f_*$ , we have

$$\begin{aligned}\beta_u(T) &= \sup_{t \in T} \left\{ f(t) + \frac{\mathbf{1}^\top \bar{f}_t''}{2\sigma u_t} + \frac{B_t}{u_t} + \mu_\sigma(t) \right\} \\ &= \sup_{t \in T} \left\{ f_*(t) + u_\tau C(t - \tau) + \frac{\mathbf{1}^\top (z_t - u_t \mu_{02} + u_\tau \mu_2(t - \tau))}{2\sigma u_t} + \frac{B_t}{u_t} + \mu_\sigma(t) \right\},\end{aligned}$$

We insert the expansions in (34), (35) and (36) into the expression of  $\beta_u(T)$  and obtain that

$$\begin{aligned}\beta_u(T) &= \sup_{t \in T} \left\{ w + y^\top(t - \tau) + \frac{1}{2}(t - \tau)^\top z(t - \tau) + g_3(t - \tau) + R_f(t - \tau) + g(t - \tau) \right. \\ &\quad + u_\tau \left( 1 - \frac{1}{2}(t - \tau)^\top(t - \tau) + C_4(t - \tau) + R_C(t - \tau) \right) \\ &\quad + \mu_\sigma(\tau) + \partial \mu_\sigma(\tau)^\top(t - \tau) + \frac{1}{2}(t - \tau)^\top \Delta \mu_\sigma(\tau)(t - \tau) + \sigma^{-1} R_\mu(t - \tau) \\ &\quad \left. + \frac{\mathbf{1}^\top (z_t - u_t \mu_{02} + u_\tau \mu_2(t - \tau))}{2\sigma u_t} + \frac{B_t}{u_t} \right\} \\ &= \sup_{t \in T} \left\{ u + w + \frac{1}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} - \frac{1}{2} (t - \tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y})^\top (uI - \tilde{\mathbf{z}}) (t - \tau - (uI - \tilde{\mathbf{z}})^{-1} \tilde{y}) \right. \\ &\quad \left. + g_3(t - \tau) + u_\tau C_4(t - \tau) + R(t - \tau) + g(t - \tau) + \frac{\mathbf{1}^\top (z_t - u_t \mu_{02} + u_\tau \mu_2(t - \tau))}{2\sigma u_t} + \frac{B_t}{u_t} \right\}.\end{aligned}$$

Note that the above display is approximately a quadratic function of  $t - \tau$  and is maximized approximately at  $t - \tau = (uI - \tilde{\mathbf{z}})^{-1} \tilde{y}$ . In addition, on the set  $\mathcal{L}$ , we have that  $|\tau - t_*| < u^{-1/2+\epsilon}$  and thus  $\tilde{y} = y + O(u^{-1/2+\epsilon})$ . Therefore, on the set  $\mathcal{L}$ , we have the following approximation of  $\beta_u(T)$

$$\mathcal{A}_\tau + \inf_{|t-\tau| < u^{-1/2+\epsilon}} g(t) \leq \beta_u(T) - u + o(u^{-1}) \leq \mathcal{A}_\tau + \sup_{|t-\tau| < u^{-1/2+\epsilon}} g(t). \quad (75)$$

Then,  $\beta_u(T) > u$  if  $\mathcal{A}_\tau + o_p(u^{-1}) > 0$ . Thus, we obtain the same representation as in Part 1 in the proof of Theorem 3.

Since we use the same change of measure, the analysis of the likelihood ratio is exactly the same as Part 2 of Theorem 3. For Part 3, we compute the second moment of  $dP/dQ$  on the set  $\{\beta_u(T) > u\}$ . This is also identical to the proof of Theorem 3. Thus, with the same choice of tuning parameters, we have that

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \beta_u(T) > u \right] \leq (1 + \varepsilon) v^2(b).$$

Additionally, Lemma 17 provides an approximation that  $P(\beta_u(T) > u) \sim v(b)$ . Thus, we use Lemma 12 (presented at the beginning of Section 3) and complete the proof.

## 5 Proof of Theorem 9

For the bias control, we need the following result that is proved in [35].

**Proposition 14** *Suppose that conditions C1-6 are satisfied. Let  $F'(x)$  be the probability density function of  $\log \mathcal{I}(T) = \log \int_T e^{\sigma f(t) + \mu(t)} dt$ . Then the following approximation holds as  $x \rightarrow \infty$*

$$F'(x) \sim \sigma^{-2} x \cdot v(e^x).$$

Thus, for any small  $\varepsilon$ ,

$$P(b < \mathcal{I}(T) < b(1 + \varepsilon/\log b) \mid \mathcal{I}(T) > b) = \Theta(\varepsilon). \quad (76)$$

Similar to the log-normal distribution, the overshoot of  $\mathcal{I}(T)$  is  $\Theta(b/\log b)$ . Note that

$$|v_M(b) - v(b)| \leq P(\mathcal{I}(T) > b, \mathcal{I}_M(T) < b) + P(\mathcal{I}(T) < b, \mathcal{I}_M(T) > b).$$

Let

$$\mathcal{L}_\varepsilon = \left\{ \sup_{t \in T} |\partial f(t)| \leq 2(1 - u^{-2} \log \varepsilon)u \right\}.$$

Note that  $\partial f(t)$  is a  $d$ -dimensional Gaussian process. Using Borel-TIS lemma, we obtain that

$$P(\mathcal{L}_\varepsilon^c) = o(1)\varepsilon \cdot v(b),$$

it is sufficient to control  $P(\mathcal{I}(T) > b, \mathcal{I}_M(T) < b, \mathcal{L}_\varepsilon)$  and  $P(\mathcal{I}(T) < b, \mathcal{I}_M(T) > b, \mathcal{L}_\varepsilon)$ .

By the definition of  $\mathcal{I}_M$  in (22), there exists a constant  $c_1 > 0$  such that

$$\begin{aligned} \Delta = |\mathcal{I}(T) - \mathcal{I}_M(T)| &\leq \sum_{i=1}^M \left| \int_{T_N(t_i)} e^{\sigma f(t) + \mu(t)} dt - \text{mes}(T_N(t_i)) \cdot e^{\sigma f(t_i) + \mu(t_i)} \right| \\ &\leq c_1 \min\{\mathcal{I}_M(T), \mathcal{I}(T)\} \cdot \sup_{t \in T} |\partial f(t)|/N \end{aligned}$$

Then we have, on the set  $\mathcal{L}_\varepsilon$ ,  $\Delta \leq 2c_1 \min\{\mathcal{I}_M(T), \mathcal{I}(T)\}(1 - u^{-2} \log \varepsilon)u/N$ , which implies that

$$P(\mathcal{I}(T) > b, \mathcal{I}_M(T) < b, \mathcal{L}_\varepsilon) \leq P(b < \mathcal{I}(T) < b(1 + 2(1 - u^{-2} \log \varepsilon)u/N)) = O(1) \frac{u(1 - u^{-2} \log \varepsilon) \log b}{N} v(b).$$

The last step is due to the result of Proposition 14 and further (76). Thus, it is sufficient to choose  $N = O(\varepsilon^{-1-\varepsilon_0} u^{2+\varepsilon_0})$  so that the above probability is bounded by  $\varepsilon v(b)$ . The bound of  $P(\mathcal{I}(T) < b, \mathcal{I}_M(T) > b, \mathcal{L}_\varepsilon)$  is completely analogous.

## 6 Proof of Theorem 10

The proof of Theorem 10 is similar to that of Theorem 3. Therefore, we only layout the key steps. The only difference is that we replace the integral by a finite sum over  $T_N$ . Recall that the proof of Theorem 3 consists of three parts: first, we write the event  $\{\mathcal{I}_M(T) > b\}$  as a function of  $(w, y, z)$  (with an ignorable correction term); second, we write the likelihood ratio as a function of  $(w, y, z)$  (with an ignorable correction term); third, we integrate the likelihood ratio with respect to  $(\iota, \tau, w, y, z)$ . For the current proof, we also have similar 3 parts.

**Part 1.** For the first step in the proof of Theorem 3, we write  $\mathcal{I}(T) > b$  if and only if  $\mathcal{A}_\tau + o(u^{-1}) > u^{-1}\sigma^{-1}\xi_u$ . With the current discretization size, as proved in Theorem 9,

$$\log \mathcal{I}(T) - \log \mathcal{I}_M(T) = o(u^{-1}).$$

Thus, we reach the same result that  $\mathcal{I}_M(T) > b$  if  $\mathcal{A}_\tau + o(u^{-1}) > u^{-1}\sigma^{-1}\xi_u$ .

**Part 2.** Consider the likelihood ratio

$$\frac{dQ}{dP} = \int_T \left[ (1 - \rho_1 - \rho_2)l(t)LR(t) + \rho_1 l(t)LR_1(t) + \frac{\rho_2}{mes(T)}LR_2(t) \right] dt.$$

Under the discretization setup, we have

$$\frac{dQ_M}{dP} = \frac{1 - \rho_1 - \rho_2}{\kappa} \sum_{i=1}^M l(t_i)LR(t_i) + \frac{\rho_1}{\kappa} \sum_{i=1}^M l(t_i)LR_1(t_i) + \rho_2 \sum_{i=1}^M \frac{1}{M} LR_2(t_i),$$

which is a discrete approximation of  $dQ/dP$ . In the proof of Theorem 3, after taking all the terms not consisting of  $t$  out of the integral (such as that in (49)), the discrete sum is essentially approximating the following integral

$$\int_{|t-\tau| < u^{-1+\delta'}} \exp \left\{ -\frac{(1-\lambda)(u_\tau + \zeta_u)}{2} (t - \tau - (uI - \tilde{\mathbf{z}})^{-1}\tilde{\mathbf{y}})^\top (uI - \tilde{\mathbf{z}}) (t - \tau - (uI - \tilde{\mathbf{z}})^{-1}\tilde{\mathbf{y}}) \right\} dt.$$

The above integral concentrates on a region of size  $O(u^{-1})$ . Given that we choose  $N > u^2$ , the discretized likelihood ratio in  $dQ_M/dP$  approximate  $dQ/dP$  up to a constant in the sense that

$$\frac{dQ_M}{dP} = \Theta(1) \frac{dQ}{dP} \tag{77}$$

**Part 3.** With the results of Parts 1 and 2, the analysis of Part 3 is completely analogous to Part 3 in the proof of Theorem 3. Thus, we conclude that

$$E^{Q_M}(\tilde{L}_b^2) \leq \kappa_1 v(b)^2,$$

where the constant  $\kappa_1$  depends on the  $\Theta(1)$  in (77).

## 7 Appendix: The Lemmas

In this section, we state all the lemmas used in the previous section. To facilitate reading, we move several lengthy proofs (Lemmas 16, 17, 19, 21, 22, and 23) to the supplemental materials, as those proofs are not particularly related to the proof of the theorems and mostly involve tedious elementary algebra.

The first lemma is known as Borel-TIS lemma, which was proved independently by [18, 47].

**Lemma 15 (Borel-TIS)** *Let  $f(t)$ ,  $t \in \mathcal{U}$ ,  $\mathcal{U}$  is a parameter set, be mean zero Gaussian random field.  $f$  is almost surely bounded on  $\mathcal{U}$ . Then,*

$$E(\sup_{\mathcal{U}} f(t)) < \infty,$$

and

$$P(\max_{t \in \mathcal{U}} f(t) - E[\max_{t \in \mathcal{U}} f(t)] \geq b) \leq e^{-\frac{b^2}{2\sigma_{\mathcal{U}}^2}},$$

where

$$\sigma_{\mathcal{U}}^2 = \max_{t \in \mathcal{U}} \text{Var}[f(t)].$$

**Lemma 16** *Conditional on the set  $\mathcal{L}$  as defined in (37), we have that*

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b, \mathcal{L}^c \right] = o(1)v^2(b).$$

**Lemma 17** *On the set  $\mathcal{L}$  as defined in (37), we have that*

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \beta_u(T) > u, \mathcal{L}^c \right] = o(1)P(\beta_u(T) > u)^2, \quad E^Q \left[ \frac{dP}{dQ} ; \beta_u(T) > u, \mathcal{L}^c \right] = o(1)P(\beta_u(T) > u).$$

In addition, we have the approximation  $P(\beta_u(T) > u) \sim v(b)$ .

**Lemma 18** *Let  $\xi_u$  be as defined in (44), then there exists a constant  $\varpi > 0$  such that for all  $x > 0$*

$$P(u^{1/2-3\delta}|\xi_u| > x) \leq e^{-\varpi x^2} + e^{-\varpi u^2},$$

for  $u$  sufficiently large.

**Proof of Lemma 18.** We split the expectation into two parts  $\{|\tilde{S}| \leq u^\delta\}$  and  $\{|\tilde{S}| > u^\delta, \tau + (uI - \mathbf{z})^{-1/2}\tilde{S} \in T\}$ . Note that  $|S| \leq \kappa u^\delta$  and  $g(t)$  is a mean zero Gaussian random field with  $\text{Var}(g(t)) = O(|t|^6)$ . A direct application of the Borel-TIS inequality (Lemma 15) yields the result of this lemma. ■

**Lemma 19** *Let  $S$  be a random variable taking values in  $\{s : (uI - \mathbf{z})^{-1/2}s + \tau \in T\}$  with density proportional to*

$$\exp \left\{ -\frac{\sigma}{2} \left( s - (uI - \mathbf{z})^{-1/2}\tilde{y} \right)^\top \left( s - (uI - \mathbf{z})^{-1/2}\tilde{y} \right) \right\}.$$

If  $|y| \leq u^{1/2+\varepsilon}$  and  $|z| \leq u^{1/2+\varepsilon}$ , then

$$\begin{aligned} & \log \left\{ E \exp \left[ \sigma g_3 \left( (uI - \mathbf{z})^{-\frac{1}{2}} S \right) + \sigma (u - \mu_\sigma(\tau)) C_4 \left( (uI - \mathbf{z})^{-\frac{1}{2}} S \right) + \sigma R \left( (uI - \mathbf{z})^{-\frac{1}{2}} S \right) \right] \right\} \\ &= -\frac{\sigma}{8u} \left( u^{-1}\tilde{Y} + \mathbf{1}/\sigma \right)^\top \mu_{22} \left( u^{-1}\tilde{Y} + \mathbf{1}/\sigma \right) + \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma u} + \frac{1}{8\sigma u} \sum_i \partial_{iii}^4 C(0) + o(u^{-1}), \end{aligned}$$

where the expectation is taken with respect to  $S$  and  $\tilde{Y} = \{\tilde{y}_i^2, i = 1, \dots, d; 2\tilde{y}_i\tilde{y}_j, 1 \leq i < j \leq d\}$ .



**Lemma 20**

$$\log(\det(I - u^{-1}\mathbf{z})) = -u^{-1}\text{Tr}(\mathbf{z}) + \frac{1}{2}u^{-2}\mathbf{I}_2(\mathbf{z}) + o(u^{-2}),$$

where  $\text{Tr}$  is the trace of a matrix,  $\mathbf{I}_2(\mathbf{z}) = \sum_{i=1}^d \lambda_i^2$ , and  $\lambda_i$ 's are the eigenvalues of  $\mathbf{z}$ .

**Proof of Lemma 20.** The result is immediate by noting that  $\det(I - u^{-1}\mathbf{z}) = \prod_{i=1}^d (1 - \lambda_i/u)$ , and  $\text{Tr}(\mathbf{z}) = \sum_{i=1}^d \lambda_i$ . ■

**Lemma 21** On the set  $\mathcal{L}$ ,  $\mathcal{I}_2$  defined as in (47) can be written as

$$\begin{aligned} & \int_{A^*, |t-\tau| < u^{-1+\delta'}} \exp \left\{ -\frac{u_{t_*}(t-t_*)^\top \Delta \mu_\sigma(t_*)(t-t_*)}{2} + \frac{u_t^2}{2} \right\} \times u_t \\ & \times \exp \left\{ (1-\lambda)u_t [w_t + u_\tau C(t-\tau) - u_t] + (1-\lambda) \frac{\mathbf{1}^\top (z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau)}{2\sigma} - \lambda B_t - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\ & \times \exp \left\{ \frac{(w_t + u_\tau C(t-\tau) - u_t)^2 - 2(w_t + u_\tau C(t-\tau) - u_t) \mu_{20} \mu_{22}^{-1} (z_t - \mu_{02}u_t + \mu_2(t-\tau)u_\tau)}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} \right\} dt. \end{aligned}$$

**Lemma 22** For  $\eta > 0$ , on the set  $\mathcal{L}$ , if  $A_\tau \geq 0$ , then

$$\{|t-\tau| \leq u^{-1+\delta'}\} \subseteq A^*.$$

**Lemma 23** Consider that  $\{t : |t-\tau| \leq u^{-1/2+\delta'}\} \subset T$ . There exists some  $\delta^* > 0$  such that for all  $-u^{3/2+\epsilon} < x < 0$ ,

$$E \left[ \frac{1}{\rho_1^2 \kappa_{1,2}^2}; x > \xi_u \middle| \imath, \tau, w, y, z, C_1 \right] = O(1) \rho_1^{-2} e^{u^{\delta^*} x},$$

and

$$E \left[ \frac{1}{(1 - \rho_1 - \rho_2)^2 \kappa_{2,1}^2}; x > \xi_u \middle| \imath, \tau, w, y, z, C_2 \right] = O(1) (1 - \rho_1 - \rho_2)^{-2} (1 - \lambda)^{-d} e^{u^{\delta^*} x},$$

where  $C_1 = \{\text{mes}(A^c \cap D) \geq \text{mes}(A \cap D)\}$ , and  $C_2 = \{\text{mes}(A^c \cap D) < \text{mes}(A \cap D)\}$ .

## References

- [1] R.J. Adler. *The Geometry of Random Fields*. Wiley, Chichester, U.K.; New York, U.S.A., 1981.
- [2] R.J. Adler, G. Samorodnitsky, and J.E. Taylor. High level excursion set geometry for non-gaussian infinitely divisible random fields. *preprint*, 2009.
- [3] R.J. Adler and J.E. Taylor. *Random fields and geometry*. Springer, 2007.
- [4] S.M. Ahsan. Portfolio selection in a lognormal securities market. *Zeitschrift Fur Nationalokonomie-Journal of Economics*, 38(1-2):105–118, 1978.
- [5] S. Asmussen. *Ruin Probabilities*. World Scientific, London, UK, 2000.
- [6] S. Asmussen and C. Klüppelberg. Large deviations results for subexponential tails, with applications to insurance risk. *Stochastic Processes and their Applications*, 64:103–125, 1996.

- [7] S. Asmussen and L. Rojas-Nandayapa. Asymptotics of sums of lognormal random variables with gaussian copula. *Statistics & Probability Letters*, 78:2709–2714, 2008.
- [8] J. M. Azais and M. Wschebor. On the distribution of the maximum of a gaussian field with  $d$  parameters. *Annals of Applied Probability*, 15(1A):254–278, 2005.
- [9] J. M. Azais and M. Wschebor. A general expression for the distribution of the maximum of a gaussian field and the approximation of the tail. *Stochastic Processes and Their Applications*, 118(7):1190–1218, 2008.
- [10] J. M. Azais and M. Wschebor. *Level sets and extrema of random processes and fields*. Wiley, Hoboken, N.J., 2009.
- [11] S. Basak and A. Shapiro. Value-at-risk-based risk management: Optimal policies and asset prices. *Review of Financial Studies*, 14(2):371–405, 2001.
- [12] S. M. Berman. An asymptotic formula for the distribution of the maximum of a gaussian process with stationary increments. *Journal of Applied Probability*, 22(2):454–460, 1985.
- [13] F. Black and M. Scholes. Pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [14] J. H. Blanchet, J. Liu, and X. Yang. Monte carlo for large credit portfolios with potentially high correlations. *Proceedings of the 2010 Winter Simulation Conference*, 2010.
- [15] J. H. Blanchet and J. C. Liu. State-dependent importance sampling for regularly varying random walks. *Advances in Applied Probability*, 40:1104–1128, 2008.
- [16] Jose Blanchet and Jingchen Liu. Efficient importance sampling in ruin problems for multidimensional regularly varying random walks. *Journal of Applied Probability*, 47:301–322, 2010.
- [17] Jose Blanchet and Jingchen Liu. Efficient simulation and conditional functional limit theorems for ruinous heavy-tailed random walks. *preprint*, <http://arxiv.org/abs/1006.2808>, 2012.
- [18] C. Borell. The Brunn-Minkowski inequality in Gauss space. *Inventiones Mathematicae*, 1975.
- [19] M. J. Campbell. Time-series regression for counts - an investigation into the relationship between sudden-infant-death-syndrome and environmental-temperature. *Journal of the Royal Statistical Society Series a-Statistics in Society*, 157:191–208, 1994.
- [20] K. S. Chan and J. Ledolter. Monte-carlo em estimation for time-series models involving counts. *Journal of the American Statistical Association*, 90(429):242–252, 1995.
- [21] D. R. Cox. Some statistical methods connected with series of events. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 17(2):129–164, 1955.
- [22] D. R. Cox and Valerie Isham. *Point processes*. Monographs on applied probability and statistics. Chapman and Hall, London ; New York, 1980.
- [23] R. A. Davis, W. T. M. Dunsmuir, and Y. Wang. On autocorrelation in a poisson regression model. *Biometrika*, 87(3):491–505, 2000.
- [24] H. P. Deutsch. *Derivatives and internal models*. Finance and capital markets. Palgrave Macmillan, Houndmills, Basingstoke, Hampshire ; New York, N.Y., 3rd edition, 2004.

- [25] D. Duffie and J. Pan. An overview of value at risk. *The Journal of Derivatives*, 4(3):7–49, 1997.
- [26] D. Dufresne. The integral of geometric brownian motion. *Advances in Applied Probability*, 33(1):223–241, 2001.
- [27] P. Dupuis and R. S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley-Interscience, 1997.
- [28] S. Foss and A. Richards. On sums of conditionally independent subexponential random variables. *Mathematics of Operations Research*, 35:102–119, 2010.
- [29] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Variance reduction techniques for estimating value-at-risk. *Management Science*, 46(10):1349–1364, 2000.
- [30] Jürg Hüsler. Extreme values and high boundary crossings of locally stationary gaussian processes. *Annals of Probability*, 18:1141–1158, 1990.
- [31] Jürg Hüsler, Vladimir Piterbarg, and Yueming Zhang. Extremes of gaussian processes with random variance. *Electronic Journal of Probability*, 16(45):1254–1280, 2011.
- [32] H. J. Landau and L. A. Shepp. Supremum of a gaussian process. *Sankhya-the Indian Journal of Statistics Series A*, 32(Dec):369–378, 1970.
- [33] M. Ledoux and M. Talagrand. Probability in banach spaces : isoperimetry and processes. 1991.
- [34] J. Liu. Tail approximations of integrals of gaussian random fields. *Annals of Probability*, to appear, available at <http://arxiv.org/abs/1006.2837>, 2011.
- [35] J. Liu and G. Xu. On the density functions of integrals of gaussian random fields. *preprint*, <http://stat.columbia.edu/~jcliu/paper/grf10.pdf>, 2012.
- [36] J. Liu and G. Xu. Some asymptotic results of gaussian random fields with varying mean functions and the associated processes. *Annals of Statistics*, 40:262–293, 2012.
- [37] M. B. Marcus and L. A. Shepp. Continuity of gaussian processes. *Transactions of the American Mathematical Society*, 151(2), 1970.
- [38] R. Merton. The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183, 1973.
- [39] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.
- [40] Y. Nardi, D. O. Siegmund, and B. Yakir. The distribution of maxima of approximately gaussian random fields. *Annals of Statistics*, 36(3):1375–1403, 2008.
- [41] V. I. Piterbarg. *Asymptotic methods in the theory of Gaussian processes and fields*. American Mathematical Society, Providence, R.I., 1996.
- [42] V.N. Sudakov and B.S. Tsirelson. Extremal properties of half spaces for spherically invariant measures. *Zap. Nauchn. Sem. LOMI*, 45:75–82, 1974.

- [43] J. Y. Sun. Tail probabilities of the maxima of gaussian random-fields. *Annals of Probability*, 21(1):34–71, 1993.
- [44] M. Talagrand. Majorizing measures: The generic chaining. *Annals of Probability*, 24(3):1049–1103, 1996.
- [45] J. Taylor, A. Takemura, and R. J. Adler. Validity of the expected euler characteristic heuristic. *Annals of Probability*, 33(4):1362–1396, 2005.
- [46] J. Traub, G. Wasilokowski, and H. Wozniakowski. *Information-Based Complexity*. Academic Press, New York, NY, 1988.
- [47] B.S. Tsirelson, I.A. Ibragimov, and V.N. Sudakov. Norms of Gaussian sample functions. *Proceedings of the Third Japan-USSR Symposium on Probability Theory (Tashkent, 1975)*, 550:20–41, 1976.
- [48] H. Wozniakowski. Computational complexity of continuous problems, 1996. Technical Report.
- [49] M. Yor. On some exponential functionals of brownian-motion. *Advances in Applied Probability*, 24(3):509–531, 1992.
- [50] S. L. Zeger. A regression-model for time-series of counts. *Biometrika*, 75(4):621–629, 1988.

## Supplementary materials: Proof of the Lemmas

**Proof of Lemma 16.** Define another set

$$\mathcal{L}_1 = \left\{ |f(\tau) - u| < u^{1/2+\epsilon}, \quad |y| < u^{1/2+\epsilon}, \quad \sup_{|t-\tau| < u^{-1/2+\delta}} |z_t| < u^{1/2+\epsilon}, \right. \\ \left. \sup_{|t-\tau| < u^{-1/2+\delta}} |\partial^3 f(t) - u_\tau \partial^3 C(t - \tau)| < u^{1/2+\epsilon}, \quad \sup_{|t-\tau| < u^{-1/2-\delta}} |g(t)| < u^{-1-\delta} \right\}. \quad (78)$$

Note that  $\mathcal{L} \subset \mathcal{L}_1$ . We first show that

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \int_T e^{\mu(t)+\sigma f(t)} dt > b, \mathcal{L}_1^c \right] = o(1)v(b)^2.$$

In this proof, we mainly use the last component of the mixture  $LR_2(t)$  that has not been used in the main proof of the theorem. Note that  $dQ/dP \geq \rho_2 \int_T LR_2(t)/mes(T)dt$  and therefore

$$\begin{aligned} & E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \int_T e^{\mu(t)+\sigma f(t)} dt > b, \mathcal{L}_1^c \right] \\ & \leq E^Q \left[ \left( \rho_2 \int_T \frac{1}{mes(T)} LR_2(t) dt \right)^{-2} ; \mathcal{I}(T) > b, \mathcal{L}_1^c \right] \\ & = \frac{1}{\rho_2^2} E^Q \left[ \left( \frac{1}{mes(T)} \int_T e^{-\frac{1}{2}u_t^2 + u_t f(t)} dt \right)^{-2} ; \frac{\mathcal{I}(T)}{mes(T)} > \frac{b}{mes(T)}, \mathcal{L}_1^c \right]. \end{aligned} \quad (79)$$

On the set

$$\frac{1}{mes(T)} \int_T e^{\sigma(\mu_\sigma(t)+f(t))} dt = \frac{\mathcal{I}(T)}{mes(T)} > \frac{b}{mes(T)}$$

we have for large  $b$ ,

$$\begin{aligned} & \frac{1}{mes(T)} \int_T e^{-\frac{1}{2}u_t^2 + u_t f(t)} dt \\ & = e^{-\frac{1}{2}u^2} \times \frac{1}{mes(T)} \int_T e^{(u-\mu_\sigma(t)) \cdot (f(t)+\mu_\sigma(t)) + \frac{1}{2}\mu_\sigma^2(t)} dt \\ & \geq e^{-\frac{1}{2}u^2 + \frac{1}{2} \min_{t \in T} \mu_\sigma^2(t)} \times \frac{1}{mes(T)} \times \left[ \int_{T \cap \{f(t)+\mu_\sigma(t) \geq 0\}} e^{(u-\max_{t \in T} \mu_\sigma(t))(f(t)+\mu_\sigma(t))} dt \right. \\ & \quad \left. + \int_{T \cap \{f(t)+\mu_\sigma(t) < 0\}} e^{(u-\min_{t \in T} \mu_\sigma(t))(f(t)+\mu_\sigma(t))} dt \right] \\ & \geq \frac{1}{2} e^{-\frac{1}{2}u^2 + \frac{1}{2} \min_{t \in T} \mu_\sigma^2(t)} \times \left[ \frac{1}{mes(T)} \int_T e^{(u-\max_{t \in T} \mu_\sigma(t))(f(t)+\mu_\sigma(t))} dt \right]. \end{aligned} \quad (80)$$

We apply Jensen's inequality to the term in the bracket and have that

$$\begin{aligned} (80) & \geq \frac{1}{2} e^{-\frac{1}{2}u^2 + \frac{1}{2} \min_{t \in T} \mu_\sigma^2(t)} \times \left[ \frac{1}{mes(T)} b \right]^{(u-\max_{t \in T} \mu_\sigma(t))/\sigma} \\ & \geq e^{\frac{1}{2}u^2 - c_1 u \log u}, \end{aligned}$$

where  $c_1$  is some constant. The last step of the above display uses the fact that  $u - \log b = O(\log u)$ . Therefore, the above derivation implies that

$$(79) \leq \rho_2^{-2} e^{-u^2 + 2c_1 u \log u} \times Q(\mathcal{L}_1^c)$$

By Borel-TIS Lemma (Lemma 15), we have that  $Q(\mathcal{L}_1^c) = O(1)e^{-\omega u^{1+2\epsilon}}$  for some  $\omega > 0$  and thus

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b, \mathcal{L}_1^c \right] = o(1) \rho_2^{-2} e^{-u^2 - \frac{\omega}{2} u^{1+2\epsilon}} = o(1) v^2(b).$$

We now proceed to bound for the rest of the expectation, that is

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b, \mathcal{L}^c \cap \mathcal{L}_1 \right].$$

This requires some fine analysis. Similar to the proof on the set  $\mathcal{L}_1$ , we still focus on the last component of the likelihood ratio corresponding to  $LR_2$ . Let

$$\mathcal{L}_2 = \mathcal{L}^c \cap \mathcal{L}_1.$$

Since  $\mathcal{L}_2 \subset \mathcal{L}_1$ , the rest of the derivations are on the set  $\mathcal{L}_1$ . Note that

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2 ; \mathcal{I}(T) > b, \mathcal{L}_2 \right] \leq \text{mes}^2(T) \rho_2^{-2} E^Q \left[ \left( \int_T e^{-\frac{1}{2} u_t^2 + [w_t + u_\tau C(t-\tau)] u_t} dt \right)^{-2} ; \mathcal{I}(T) > b, \mathcal{L}_2 \right].$$

The derivation for the above expectation takes a similar three-step procedure as that in the proof of Theorem 3. We only state the key steps here.

**Part 1.** On the set  $\mathcal{L}_1$  and for a given  $\tau$ , following Part 1 in the proof of Theorem 3, in particular results in (45), we have  $\mathcal{I}(T) > b$  if and only if

$$\mathcal{A}' + o(u^{-1}) > u^{-1} \sigma^{-1} \xi_u,$$

where

$$\begin{aligned} \mathcal{A}' &= w + \frac{1}{2} \tilde{y}^\top (uI - \tilde{\mathbf{z}})^{-1} \tilde{y} - \frac{1}{2\sigma} \log \det(I - u^{-1} \tilde{\mathbf{z}}) - \frac{1}{8u} (u^{-1} \tilde{Y} + \mathbf{1}/\sigma)^\top \mu_{22} (u^{-1} \tilde{Y} + \mathbf{1}/\sigma) \\ &\quad + \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2 u} + \frac{1}{8\sigma^2 u} \sum_i \partial_{iiii}^4 C(0). \end{aligned}$$

**Part 2.** Let  $\lambda_u = u^{-1/2+\delta}$  and we have that

$$\begin{aligned} & \int_T e^{-\frac{1}{2} u_t^2 + [w_t + u_\tau C(t-\tau)] u_t} dt \\ & \geq e^{-u^2/2} \int_{|t-\tau| < \lambda_u} e^{u_t [w_t + u_\tau C(t-\tau) + \mu_\sigma(t)] + \frac{1}{2} \mu_\sigma^2(t)} dt \\ & = (1 + o(1)) e^{u^2/2 - u\mu_\sigma(\tau) + \mu_\sigma^2(\tau)/2} \int_{|t-\tau| < \lambda_u} e^{u_t [w_t + u_\tau C(t-\tau) - u + \mu_\sigma(\tau)]} dt. \end{aligned}$$

Similar to Part 1, we insert the Taylor expansion of  $w_t$  and  $C(t - \tau)$  to the above display and the integral equals ( $\zeta_u = O(u^{-1/2+\delta})$ )

$$\begin{aligned}
& \int_{|t-\tau| < \lambda_u} e^{u_t[w_t+u_\tau C(t-\tau)-u+\mu_\sigma(\tau)]} dt \\
&= (1+o(1)) \exp \left\{ (u_\tau + \zeta_u) \left[ w + \frac{1}{2} y(u_\tau I - \mathbf{z})^{-1} y \right] \right\} \\
& \quad \times \int_{|t-\tau| < \lambda_u} \exp \left\{ (u_\tau + \zeta_u) \left[ -\frac{1}{2} (t - \tau - (u_\tau I - \mathbf{z})^{-1} y)^\top (u_\tau I - \mathbf{z}) (t - \tau - (u_\tau I - \mathbf{z})^{-1} y) \right. \right. \\
& \quad \left. \left. + g_3(t - \tau) + u_\tau C_4(t - \tau) + g(t - \tau) \right] \right\} dt.
\end{aligned} \tag{81}$$

We further restrict the interval to the region  $S_M = \{t : |t - \tau - (u_\tau I - \mathbf{z})^{-1} y| \leq M u^{-1}\}$  for some  $M$  large. On the set  $S_M$ ,  $u g(t - \tau) = o(1)$ . Then,

$$\begin{aligned}
& \int_{|t-\tau| < \lambda_u} e^{u_t[w_t+u_\tau C(t-\tau)-u+\mu_\sigma(\tau)]} dt \\
& \geq (1+o(1)) \exp \left\{ (u_\tau + \zeta_u) \left[ w + \frac{1}{2} y(u_\tau I - \mathbf{z})^{-1} y \right] + o(|y|^2) \right\} \\
& \quad \times \int_{S_M} \exp \left\{ (u_\tau + \zeta_u) \left[ -\frac{1}{2} (t - \tau - (u_\tau I - \mathbf{z})^{-1} y)^\top (u_\tau I - \mathbf{z}) (t - \tau - (u_\tau I - \mathbf{z})^{-1} y) \right] \right\} dt \\
& \geq \delta_0 u^{-d} \exp \left\{ (u_\tau + \zeta_u) \left[ w + \frac{1}{2} y(u_\tau I - \mathbf{z})^{-1} y \right] + o(|y|^2) \right\}.
\end{aligned}$$

We insert the above result back to (81) and obtain that

$$\begin{aligned}
& \int_T e^{-\frac{1}{2} u_t^2 + [w_t + u_\tau C(t-\tau)] u_t} dt \\
& \geq \delta_0^{-1} u^{-d} e^{u_\tau^2/2} \exp \left\{ (u_\tau + \zeta_u) \left[ w + \frac{1}{2} y(u_\tau I - \mathbf{z})^{-1} y \right] + o(|y|^2) \right\} \\
& = \Theta(1) u^{-d} e^{u_\tau^2/2} \exp \left\{ (u_\tau + \zeta_u) \mathcal{A}' + o(|y|^2) + O(|z|) \right\}.
\end{aligned}$$

On the set  $\mathcal{L}_1$ , note that

$$\mathcal{A}' = \mathcal{A}_\tau + o(u^{-1}|y|^2) + O(u^{-1} + u^{-1}|z|).$$

Then, we have that

$$\int_T e^{-\frac{1}{2} u_t^2 + [w_t + u_\tau C(t-\tau)] u_t} dt = \Theta(1) u^{-d} e^{u_\tau^2/2} \exp \left\{ (u_\tau + \zeta_u) \mathcal{A}_\tau + o(|y|^2) + O(|z|) \right\}$$

**Part 3.** With the results of Parts 1 and 2, the analysis of this part is completely analogous to Part 3 of the proof of Theorem 3. For  $i = 0$  we have that

$$\begin{aligned}
& E^Q \left[ \left( \int_T e^{-\frac{1}{2} u_t^2 + [w_t + u_\tau C(t-\tau)] u_t} dt \right)^{-2}; \mathcal{I}(T) > b, \mathcal{L}_2, i = 0 \right] \\
&= O(1) u^{2d} E^Q \left\{ e^{-u_\tau^2} E_{i,\tau}^Q \left[ e^{-2(u_\tau + \zeta_u) \mathcal{A}_\tau + o(|y|^2) + O(|z|)}; \mathcal{I}(T) > b \right]; |\tau - t_*| > u^{-1/2+\epsilon}, i = 0 \right\} \\
& \quad + O(1) u^{2d} E^Q \left\{ e^{-u_\tau^2} E_{i,\tau}^Q \left[ e^{-2(u_\tau + \zeta_u) \mathcal{A}_\tau + o(|y|^2) + O(|z|)}; \mathcal{I}(T) > b, \mathcal{L}_2 \right]; |\tau - t_*| \leq u^{-1/2+\epsilon}, i = 0 \right\}.
\end{aligned}$$

The first term is on the set  $|\tau - t_*| > u^{-1/2+\epsilon}$  and therefore  $e^{-u_\tau^2} \leq e^{-u_{t_*}^2 - u^\epsilon}$ . For the second term, on the set  $|\tau - t_*| \leq u^{-1/2+\epsilon}$ , we have  $|y| > u^\epsilon$  or  $\sup |z_t| > u^\epsilon$ , the inner expectation is

$$E_{\iota, \tau}^Q \left[ e^{-2(u_\tau + \zeta_u) \mathcal{A}_\tau + o(|y|^2) + O(|z|)}; \mathcal{I}(T) > b, \mathcal{L}_2 \right] = O(1) e^{-u^{2\epsilon}/2}.$$

For the case that  $\iota = 1$ , the handling is similar. Therefore, the overall contribution is

$$E^Q \left[ \left( \int_T e^{-\frac{1}{2}u_t^2 + [w_t + u_\tau C(t-\tau)]u_t} dt \right)^{-2}; \mathcal{I}(T) > b, \mathcal{L}_2 \right] = o(1)v^2(b).$$

Thereby, we conclude the proof. ■

**Proof of Lemma 17.** The proof of Lemma 17 is analogous to that of Lemma 16. We only show the key steps here.

Consider the set

$$\mathcal{L}_3 = \left\{ \sup_{t \in T} |f(t)| < Mu, \quad \sup_{t \in T} |\partial f(t)| < M^2 u, \quad \sup_{t \in T} |\partial^2 f(t)| < M^2 u \right\},$$

where  $M$  is some big constant. By the Borel-TIS lemma, we have

$$E^Q \left[ \frac{dP}{dQ}; \beta_u(T) > u, \mathcal{L}_3^c \right] = P(\beta_u(T) > u, \mathcal{L}_3^c) \leq P(\mathcal{L}_3^c) = o(1)v(b).$$

For the second moment, we have that

$$\begin{aligned} E^Q \left[ \left( \frac{dP}{dQ} \right)^2; \beta_u(T) > u, \mathcal{L}_3^c \right] &= E \left[ \frac{dP}{dQ}; \beta_u(T) > u, \mathcal{L}_3^c \right] \\ &\leq E \left[ \left( \int_T e^{-\frac{1}{2}u_t^2 + u_t f(t)} dt \right)^{-1}; \mathcal{L}_3^c \right] \\ &\leq E \left[ e^{u^2 + u \sup |f(t)|}; \mathcal{L}_3^c \right]. \end{aligned}$$

We can always choose  $M$  large enough such that the term is of order  $o(v^2(b))$ . On the set  $\mathcal{L}_3$ , let  $t_{sup}$  be the maximum of  $f(t)$  that is  $f(t_{sup}) = \sup_{t \in T} f(t)$ . Then, there exists constant  $c_1$  such that

$$\begin{aligned} \int_T LR_2(t) dt &= \int_T e^{-\frac{1}{2}u_t^2 + u_t f(t)} dt \\ &\geq \int_{|t - t_{sup}| < u^{-2-\epsilon}} e^{-\frac{1}{2}u_t^2 + u_t f(t)} dt \\ &\geq e^{-\frac{u^2 + c_1 u}{2} + u f(t_{sup})}. \end{aligned} \tag{82}$$

Now consider the set  $\mathcal{L}_1$  defined in (78), we have

$$\begin{aligned} &E^Q \left[ \left( \frac{dP}{dQ} \right)^2; \beta_u(T) > u, \mathcal{L}_1^c \cap \mathcal{L}_3 \right] \\ &\leq E^Q \left[ \left( \rho_2 \int_T \frac{1}{mes(T)} LR_2(t) dt \right)^{-2}; \beta_u(T) > u, \mathcal{L}_1^c \cap \mathcal{L}_3 \right] \\ &= \frac{mes^2(T)}{\rho_2^2} E^Q \left[ \left( \int_T e^{-\frac{1}{2}u_t^2 + u_t f(t)} dt \right)^{-2}; \beta_u(T) > u, \mathcal{L}_1^c \cap \mathcal{L}_3 \right]. \end{aligned} \tag{83}$$



Then, on the set  $\mathcal{L}_3$ , there exists a constant  $c_2$  such that

$$(83) \leq \frac{mes^2(T)}{\rho_2^2} E^Q \left[ \left( \int_T e^{-\frac{1}{2}u_t^2 + u_t f(t)} dt \right)^{-2}; f(t_{sup}) > u - c_2, \mathcal{L}_1^c \cap \mathcal{L}_3 \right] \\ \leq \frac{mes^2(T)}{\rho_2^2} E^Q \left[ \left( e^{-\frac{u^2 + c_1 u}{2} + u f(t_{sup})} \right)^{-2}; f(t_{sup}) > u - c_2, \mathcal{L}_1^c \cap \mathcal{L}_3 \right],$$

where in the last step we plugged in the bound in (82). Therefore there exists a constant  $c_3$  such that

$$(83) \leq \rho_2^{-2} e^{-u^2 + 2c_3 u \log u} Q(\mathcal{L}_1^c \cap \mathcal{L}_3).$$

Then the Borel-TIS Lemma implies that on the set  $\mathcal{L}_1^c$ , there exists a positive constant  $\varpi$  such that

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2; \beta_u(T) > u, \mathcal{L}_1^c \right] = o(1) \rho_2^{-2} e^{-u^2 - u^{1+\varpi}} = o(1) v^2(b).$$

With exactly the same argument, we have that

$$E^Q \left[ \frac{dP}{dQ}; \beta_u(T) > u, \mathcal{L}_1^c \right] = o(1) v(b)$$

Now on  $\mathcal{L}_1$ , with a similar three-step procedure explored in the proof of Lemma 16, we can obtain that

$$E^Q \left[ \left( \frac{dP}{dQ} \right)^2; \beta_u(T) > u, \mathcal{L}^c \right] = o(1) v^2(b), \quad E^Q \left[ \frac{dP}{dQ}; \beta_u(T) > u, \mathcal{L}^c \right] = o(1) v(b).$$

With a similar derivation as in Theorem 3, we obtain that

$$P(\beta_u(T) > u) \sim v(b).$$

The detailed derivation of the above asymptotic approximation is omitted. ■

**Proof of Lemma 19.** For  $\delta > \epsilon$ , we first split the expectation into two parts:

$$E \left[ \exp \left\{ \sigma g_3 \left( (uI - \mathbf{z})^{-\frac{1}{2}} S \right) + \sigma(u - \mu_\sigma(\tau)) C_4 \left( (uI - \mathbf{z})^{-\frac{1}{2}} S \right) + \sigma R \left( (uI - \mathbf{z})^{-\frac{1}{2}} S \right) \right\} \right] \\ = E \left[ \dots; |S| \leq u^\delta \right] + E \left[ \dots; |S| > u^\delta \right] \\ = J_1 + J_2.$$

Let  $t = (uI - \mathbf{z})^{-\frac{1}{2}} s$ . Given that  $C(t)$  is a monotone non-increasing function, we have that for all  $|s| = O(u^\delta)$

$$-\frac{\sigma}{2} \left( s - (uI - \mathbf{z})^{-1/2} \tilde{y} \right)^\top \left( s - (uI - \mathbf{z})^{-1/2} \tilde{y} \right) \\ + \sigma g_3 \left( (uI - \mathbf{z})^{-\frac{1}{2}} s \right) + \sigma(u - \mu_\sigma(\tau)) C_4 \left( (uI - \mathbf{z})^{-\frac{1}{2}} s \right) + \sigma R \left( (uI - \mathbf{z})^{-\frac{1}{2}} s \right) \\ = -\frac{\sigma}{2} \left( t - (uI - \mathbf{z})^{-1/2} \tilde{y} \right)^\top (uI - \mathbf{z}) \left( t - (uI - \mathbf{z})^{-1/2} \tilde{y} \right) \\ + \sigma g_3(t) + \sigma(u - \mu_\sigma(\tau)) C_4(t) + \sigma R(t) \\ \leq -\lambda u |t|^2 = -\lambda s^2.$$

Therefore, for  $\lambda'$  small

$$J_2 = O\left(e^{-\lambda' u^{2\delta}}\right).$$

We now consider the leading term  $J_1$ . On the set that  $|S| \leq u^\delta$  and  $|y| \leq u^{1/2+\epsilon}$ , we have that

$$\sigma g_3\left((uI - \mathbf{z})^{-\frac{1}{2}}S\right) + \sigma(u - \mu_\sigma(\tau))C_4\left((uI - \mathbf{z})^{-\frac{1}{2}}S\right) + \sigma R\left((uI - \mathbf{z})^{-\frac{1}{2}}S\right) = O\left(u^{-1+4\delta}\right).$$

By using Taylor's expansion twice, we can essentially move the expectation into the exponent and obtain that  $\log J_1$  equals

$$E\left[\sigma g_3\left((uI - \mathbf{z})^{-\frac{1}{2}}S\right) + \sigma(u - \mu_\sigma(\tau))C_4\left((uI - \mathbf{z})^{-\frac{1}{2}}S\right) + \sigma R\left((uI - \mathbf{z})^{-\frac{1}{2}}S\right); |S| \leq u^\delta\right] + o(u^{-1}).$$

Note that  $(uI - \mathbf{z})^{-1/2}\tilde{y} = u^{-1/2}\tilde{y} + O(u^{-3/2}|\tilde{y}||z|)$  and  $y = \tilde{y} + O(1)$ . Let  $Z = (Z_1, \dots, Z_d)$  be a multivariate Gaussian random vector with mean zero and covariance function  $\sigma^{-1}I$ . Then,  $S$  is equal in distribution to  $Z + u^{-1/2}\tilde{y} + O(u^{-3/2}|\tilde{y}||z|)$ . Therefore, we obtain that

$$\begin{aligned} & \log J_1 \\ &= -\frac{\sigma + O(u^{-1+\epsilon})}{6} u^{-3/2} \sum_{ijkl} \partial_{ijkl}^4 C(0) E\left[(u^{-1/2}\tilde{y}_i + Z_i)(u^{-1/2}\tilde{y}_j + Z_j)(u^{-1/2}\tilde{y}_k + Z_k)\right] y_l \\ & \quad + \frac{\sigma + O(u^{-1+\epsilon})}{24} u^{-1} \sum_{ijkl} \partial_{ijkl}^4 C(0) E\left[(u^{-1/2}\tilde{y}_i + Z_i)(u^{-1/2}\tilde{y}_j + Z_j)(u^{-1/2}\tilde{y}_k + Z_k)(u^{-1/2}\tilde{y}_l + Z_l)\right] \\ & \quad + o(u^{-1}) \\ &= -\frac{\sigma}{6} u^{-3/2} \sum_{ijkl} \partial_{ijkl}^4 C(0) E\left[(u^{-1/2}\tilde{y}_i + Z_i)(u^{-1/2}\tilde{y}_j + Z_j)(u^{-1/2}\tilde{y}_k + Z_k)\right] y_l \\ & \quad + \frac{\sigma}{24} u^{-1} \sum_{ijkl} \partial_{ijkl}^4 C(0) E\left[(u^{-1/2}\tilde{y}_i + Z_i)(u^{-1/2}\tilde{y}_j + Z_j)(u^{-1/2}\tilde{y}_k + Z_k)(u^{-1/2}\tilde{y}_l + Z_l)\right] + o(u^{-1}), \end{aligned}$$

where the expectations are taken with respect to  $Z$ . Then

$$\begin{aligned} & \log J_1 \\ &= -\frac{\sigma}{8u^3} \sum_{ijkl} \partial_{ijkl}^4 C(0) \tilde{y}_i \tilde{y}_j \tilde{y}_k \tilde{y}_l - \frac{1}{6} u^{-3/2} \sum_{iikl} 3\partial_{iikl}^4 C(0) u^{-1/2} \tilde{y}_k \tilde{y}_l \\ & \quad + \frac{1}{24} u^{-1} \sum_{iikl} 6\partial_{iikl}^4 C(0) u^{-1} \tilde{y}_k \tilde{y}_l + \frac{1}{8\sigma u} \sum_{iiii} \partial_{iiii}^4 C(0) + o(u^{-1}) \\ &= -\frac{\sigma}{8u^3} \sum_{ijkl} \partial_{ijkl}^4 C(0) \tilde{y}_i \tilde{y}_j \tilde{y}_k \tilde{y}_l - \frac{1}{4u^2} \sum_{iikl} \partial_{iikl}^4 C(0) u^{-1/2} \tilde{y}_k \tilde{y}_l + \frac{1}{8\sigma u} \sum_{iiii} \partial_{iiii}^4 C(0) + o(u^{-1}) \\ &= -\frac{\sigma}{8u^3} \tilde{Y}^\top \mu_{22} \tilde{Y} - \frac{1}{4u^2} \tilde{Y}^\top \mu_{22} \mathbf{1} + \frac{1}{8\sigma u} \sum_{iiii} \partial_{iiii}^4 C(0) \\ &= -\frac{\sigma}{8u} (u^{-1} \tilde{Y} - \mathbf{1}/\sigma)^\top \mu_{22} (u^{-1} \tilde{Y} - \mathbf{1}/\sigma) + \frac{1}{8\sigma u} \mathbf{1}^\top \mu_{22} \mathbf{1} + \frac{1}{8\sigma u} \sum_{iiii} \partial_{iiii}^4 C(0) + o(u^{-1}). \end{aligned}$$

■

**Proof of Lemma 21.** This proof only consists of elementary algebra. We expand the exponent in the second row and the second term in the third row of (47). Furthermore, we move the first

term in the third row to the last row and move the last term in the fourth row up to the third row. Then, we obtain that

$$\begin{aligned}
\mathcal{I}_2 &= \int_{A^*, |t-\tau| < u^{-1+\delta'}} \exp \left\{ -\frac{u_{t_*}(t-t_*)^\top \Delta \mu_\sigma(t_*)(t-t_*)}{2} \right\} \\
&\times u_t \times \exp \left\{ -\lambda u_t w_t + \lambda u_t (u_t - u_\tau C(t-\tau)) + (1-\lambda) \frac{\mathbf{1}^\top (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)}{2\sigma} - \lambda B_t - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\
&\times \exp \left\{ \frac{-(z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)^\top \mu_{22}^{-1} (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)}{2} \right. \\
&\quad \left. + \frac{(z_t + \mu_2(t-\tau) u_\tau)^\top \mu_{22}^{-1} (z_t + \mu_2(t-\tau) u_\tau)}{2} \right\} \\
&\times \exp \left\{ \frac{(w_t + u_\tau C(t-\tau) - \mu_{20} \mu_{22}^{-1} (z_t + \mu_2(t-\tau) u_\tau))^2}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} - \frac{|\mu_{20} \mu_{22}^{-1} (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)|^2}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} \right\} dt.
\end{aligned}$$

We now work on the exponents in the third and the last row. In particular, we have that

$$\begin{aligned}
&\frac{-(z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)^\top \mu_{22}^{-1} (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)}{2} + \frac{(z_t + \mu_2(t-\tau) u_\tau)^\top \mu_{22}^{-1} (z_t + \mu_2(t-\tau) u_\tau)}{2} \\
&+ \frac{(w_t + u_\tau C(t-\tau) - \mu_{20} \mu_{22}^{-1} (z_t + \mu_2(t-\tau) u_\tau))^2}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} - \frac{|\mu_{20} \mu_{22}^{-1} (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)|^2}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} \\
&= \frac{\mu_{20} \mu_{22}^{-1} \mu_{02} u_t^2}{2} + u_t \mu_{20} \mu_{22}^{-1} (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau) \\
&\quad + \frac{(w_t + u_\tau C(t-\tau) - \mu_{20} \mu_{22}^{-1} \mu_{02} u_t)^2}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} \\
&\quad - \frac{2(w_t + u_\tau C(t-\tau) - \mu_{20} \mu_{22}^{-1} \mu_{02} u_t) \mu_{20} \mu_{22}^{-1} (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})}.
\end{aligned}$$

We further expand the last two rows and obtain that

$$\begin{aligned}
&= \frac{1}{2} u_t^2 + u_t w_t + u_t (u_\tau C(t-\tau) - u_t) \\
&\quad + \frac{(w_t + u_\tau C(t-\tau) - u_t)^2 - 2(w_t + u_\tau C(t-\tau) - u_t) \mu_{20} \mu_{22}^{-1} (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})}.
\end{aligned}$$

We put it back to the expression of  $\mathcal{I}_2$  and obtain that

$$\begin{aligned}
\mathcal{I}_2 &= \int_{A^*, |t-\tau| < u^{-1+\delta'}} \exp \left\{ -\frac{u_{t_*}(t-t_*)^\top \Delta \mu_\sigma(t_*)(t-t_*)}{2} + \frac{u_t^2}{2} \right\} \times u_t \\
&\times \exp \left\{ (1-\lambda) u_t [w_t + u_\tau C(t-\tau) - u_t] + (1-\lambda) \frac{\mathbf{1}^\top (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)}{2\sigma} - \lambda B_t - \frac{\mathbf{1}^\top \mu_{22} \mathbf{1}}{8\sigma^2} \right\} \\
&\times \exp \left\{ \frac{(w_t + u_\tau C(t-\tau) - u_t)^2 - 2(w_t + u_\tau C(t-\tau) - u_t) \mu_{20} \mu_{22}^{-1} (z_t - \mu_{02} u_t + \mu_2(t-\tau) u_\tau)}{2(1 - \mu_{20} \mu_{22}^{-1} \mu_{02})} \right\} dt.
\end{aligned}$$

■

**Proof of Lemma 22.** We now consider a particular  $t \in A^* \cap \{t : |t-\tau| < u^{-1+\delta'}\}$ . On the localization set  $\mathcal{L}$  and  $|t-\tau| < u^{-1+\delta'}$ , we have that

$$|z_t - z| = O(u^{-1/2+\delta'+\epsilon}), \quad |z_t| = O(u^\epsilon), \quad |y_t - y| = O(u^{-1+\delta'+\epsilon}). \quad (84)$$

In what follows, we show that, for all  $t \in \{t : |t - \tau| < u^{-1+\delta'}\}$ ,  $\mathcal{A}_\tau > 0$  implies  $t \in A^*$ . On the set  $\{t : |t - \tau| < u^{-1+\delta'}\}$  and the set  $\mathcal{L}$ , by definition,  $t \in A^*$  if

$$w_t + u_\tau C(t - \tau) - u_t + \frac{|y_t - u_\tau(t - \tau)|^2}{2u_t} + \frac{\mathbf{1}^\top(z_t - \mu_{02}u_t + \mu_2(t)u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} > -\eta u_t^{-1} + o(u^{-1}).$$

We insert the expansion of  $f_*(t)$  into the above inequality and obtain that

$$\begin{aligned} & w + y^\top(t - \tau) + \frac{1}{2}(t - \tau)^\top \mathbf{z}(t - \tau) + g_3(t - \tau) + R_f(t - \tau) + g(t - \tau) \\ & - \frac{u_\tau}{2}|t - \tau|^2 + u_\tau C_4(t - \tau) + u_\tau R_C(t - \tau) - \mu_\sigma(\tau) + \mu_\sigma(t) + \frac{|y_t|^2 - 2y_t u_\tau(t - \tau) + u_\tau^2|t - \tau|^2}{2u_t} \\ & + \frac{\mathbf{1}^\top(z_t - \mu_{02}u_t + \mu_2(t)u_\tau)}{2\sigma u_t} + \frac{B_t}{u_t} > -\eta u_t^{-1} + o(u^{-1}). \end{aligned}$$

We further insert  $\mathcal{A}_\tau$  in and obtain that

$$\begin{aligned} & \mathcal{A}_\tau + \left(y - \frac{u_\tau y_t}{u_t}\right)^\top (t - \tau) + \left(\frac{u_\tau^2}{2u_t} - \frac{u_\tau}{2}\right)|t - \tau|^2 + \frac{|y_t|^2}{2u_t} - \frac{|y|^2}{2u_\tau} \\ & + \frac{B_t}{u_t} - \frac{B_\tau}{u_\tau} + \frac{\mathbf{1}^\top(z_t - \mu_{02}u_t + \mu_2(t)u_\tau)}{2\sigma u_t} - \frac{\mathbf{1}^\top z}{2\sigma u_\tau} - \mu_\sigma(\tau) + \mu_\sigma(t) \\ & + \frac{1}{2}(t - \tau)^\top \mathbf{z}(t - \tau) + g_3(t - \tau) + R_f(t - \tau) + g(t - \tau) + u_\tau C_4(t - \tau) + u_\tau R_C(t - \tau) \\ & > -\eta u_t^{-1} \end{aligned}$$

Many terms on the left-hand-side of the above inequality are in fact  $O(u^{-1-\delta'+\epsilon})$ . Thus, we have that  $t \in A^*$  if

$$\mathcal{A}_\tau + \eta u^{-1} > -g(t - \tau) + o(u^{-1}).$$

In addition,  $\sup_{|t - \tau| < u^{-1+\delta'}} |g(t)| < u^{-1-\delta}$ . Thus, for all  $\{t : |t - \tau| < u^{-1+\delta'}\}$ ,

$$t \in A^* \Leftrightarrow \mathcal{A}_\tau > -\eta u_t^{-1} + o(u^{-1}).$$

Note that  $\eta > 0$  is fixed. Thus, we have

$$\mathcal{A}_\tau > 0 \Rightarrow \{t : |t - \tau| < u^{-1+\delta'}\} \subset A^*.$$

■

**Proof of Lemma 23.** To simplify the notation, let  $\tau = 0$ . For other values of  $\tau$ , the derivation is completely analogous. We use the notation  $E_{w,y,z,C_i}[\cdot]$  to denote  $E[\cdot|w, y, z, C_i]$  and  $P_{w,y,z}$  to denote the conditional probability given  $w, y, z$ .

Note that

$$e^{-\xi_u/u} = E\left[e^{g((uI - \mathbf{Z})^{-1/2}S)}\right],$$

where  $S$  is a random variable with density proportional to (42).

The statement of the lemma considers  $x < 0$ . In the proof, we slip the sign and consider  $0 < x < u^{3/2+\epsilon}$ . Thus, we have that

$$\{-x > \xi_u\} = \{A_1 + A_2 > e^{x/u}\},$$

where

$$\begin{aligned} A_1 &= E \left[ e^{g((uI - \mathbf{z})^{-1/2} S)}; |S| \leq u^\delta \right], \\ A_2 &= E \left[ e^{g((uI - \mathbf{z})^{-1/2} S)}; |S| > u^\delta, (uI - \mathbf{z})^{-1/2} S \in T \right]. \end{aligned}$$

Furthermore, we have that

$$\log A_1 \leq \sup_{|s| \leq u^{-1/2+\delta}} g(s),$$

and by the Borell-TIS inequality (Lemma 15), we have that

$$P \left( u \cdot \sup_{|s| \leq u^{-1/2+\delta}} g(s) > x \right) \leq e^{-\lambda'' u^{1-6\delta} x^2}.$$

In addition, on the set  $\mathcal{L}$ , the process  $g$  is localized and

$$\kappa_{1,2}^{-2} = e^{O(u \cdot \sup_{|s| \leq u^{-1/2-\delta}} |g(s)|)} = e^{O(u^{-\delta})}. \quad (85)$$

Therefore, for some  $\delta^*$  small enough, we have that

$$\begin{aligned} & E_{w,y,z,C_1} \left[ \kappa_{1,2}^{-2}; A_1 + A_2 > e^{x/u}, 0 < A_2 < e^{-\lambda u^{2\delta}} \right] \\ & \leq E_{w,y,z,C_1} \left[ e^{O(u \cdot \sup_{|s| \leq u^{-1/2-\delta}} |g(s)|)}; u \cdot \sup_{|s| \leq u^{-1/2+\delta}} g(s) + O(u \cdot e^{-\lambda u^{2\delta}}) > x \right] \\ & \leq e^{-\delta^* u^{1-6\delta} x^2}. \end{aligned}$$

Similarly, since

$$\kappa_{2,1}^{-2} = \Theta(1)(1-\lambda)^{-d} e^{O(u \cdot \sup_{|s| \leq u^{-1/2-\delta}} |g(s)|)} = \Theta(1)(1-\lambda)^{-d} e^{O(u^{-\delta})}, \quad (86)$$

we have

$$E_{w,y,z,C_2} \left[ \kappa_{2,1}^{-2}; A_1 + A_2 > e^{x/u}, 0 < A_2 < e^{-\lambda u^{2\delta}} \right] = O(1)(1-\lambda)^{-d} e^{-\delta^* u^{1-6\delta} x^2}.$$

For the remainder terms, consider  $C_1$  first

$$\begin{aligned} & E_{w,y,z,C_1} \left[ \kappa_{1,2}^{-2}; A_1 + A_2 > e^{x/u}, A_2 \geq e^{-\lambda u^{2\delta}} \right] \\ & \leq E_{w,y,z,C_1} \left[ \kappa_{1,2}^{-2}; A_2 \geq e^{-\lambda u^{2\delta}} \right] \\ & \leq E_{w,y,z,C_1} \left[ \kappa_{1,2}^{-2}; \sup_{u^\delta \leq s \leq \sqrt{u}} g(s/\sqrt{u}) - \frac{1}{2} s^2 > -\lambda u^{2\delta} \right] \end{aligned} \quad (87)$$

Since for  $\lambda$  and  $\lambda'$  sufficiently small

$$\begin{aligned} P_{w,y,z} \left( A_2 > e^{-\lambda u^{2\delta}} \right) & \leq P_{w,y,z} \left( \sup_{u^\delta \leq s \leq \sqrt{u}} g(s/\sqrt{u}) - \frac{1}{2} s^2 > -\lambda u^{2\delta} \right) \\ & \leq e^{-\lambda' u^2} \leq e^{-\lambda' u^{1/2-\epsilon} x}, \end{aligned}$$

for all  $0 < x < u^{3/2+\epsilon}$ , therefore

$$\begin{aligned}
(87) &\leq e^{-\lambda' u^2} E_{w,y,z,C_1} \left[ \kappa_{1,2}^{-2} \left| \sup_{u^\delta \leq s \leq \sqrt{u}} g(s/\sqrt{u}) - \frac{1}{2} s^2 > -\lambda u^{2\delta} \right. \right] \\
&= e^{-\lambda' u^2} E_{w,y,z,C_1} \left[ e^{O\left(u \cdot \sup_{|s| \leq u^{-1/2-\delta}} |g(s)|\right)} \left| \sup_{s \in T, |s| > u^{-1/2+\delta}} g(s) - \frac{u}{2} s^2 > -\lambda u^{2\delta} \right. \right] \\
&= e^{-\lambda' u^2} e^{O(u^{-\delta})} \leq e^{-u^{\delta^*} x}.
\end{aligned}$$

Similarly, for  $C_2$  is true, by (86), the following bound holds.

$$\begin{aligned}
&E_{w,y,z,C_2} \left[ \kappa_{2,1}^{-2}; A_1 + A_2 > e^{x/u}, A_2 \geq e^{-\lambda u^{2\delta}} \right] \\
&\leq e^{-\lambda' u^2} E_{w,y,z,C_2} \left[ \kappa_{2,1}^{-2} \left| \sup_{s \in T, |s| > u^{-1/2+\delta}} g(s) - \frac{u}{2} s^2 > -\lambda u^{2\delta} \right. \right] \\
&= O(1)(1-\lambda)^{-d} e^{-\lambda' u^2} = O(1)(1-\lambda)^{-d} e^{-u^{\delta^*} x}.
\end{aligned}$$

Combining the above results, we have the conclusion. ■